

RUSSIAN ACADEMY OF SCIENCES
THE INSTITUTE OF ORIENTAL STUDIES
ST.PETERSBURG BRANCH



Manuscripta Orientalia

International Journal for Oriental Manuscript Research

Vol. 1 No. 3 December 1995



TSESA
St. Petersburg-Helsinki

CONTENTS

<i>TEXTS AND MANUSCRIPTS: DESCRIPTION AND RESEARCH</i>	3
L. Menshikov. A Fragment of an Unknown <i>Leishu</i> from Tunhuang	3
T. Sultanov. The Structure of Islamic History Book (The Method of Analysis) .	16
<i>TO THE HISTORY OF ORIENTAL TEXTOLOGY.</i>	22
K. Kepping. The Official Name of the Tangut Empire as Reflected in the Native Tangut Texts .	22
<i>PRESENTING THE COLLECTIONS.</i>	33
T. Pang. Rare Manchu Manuscripts from the Collection of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences	33
<i>ORIENTAL MANUSCRIPTS AND NEW INFORMATION TECHNOLOGIES</i> <i>Correspondence Round table: Arabic/Farsi OCR .</i>	47
A. Matveev. Sakhr Bilingual OCR (Al-Qari' al-Ali). A User's Initial Impressions .	48
J. Bell & P. Zemanek. Test of Two Arabic OCR Programs	55
P. Roochnik. Itisalal OCR Discussion	58
<i>PRESENTING THE MANUSCRIPT</i>	63
O. Akimushkin. <i>Muraqqa'</i> . Album of the Indian and Persian Miniatures of the 16—18th Centuries and the Models of the Persian Calligraphy of the Same Period	63
<i>BOOK AND SOFTWARE REVIEW</i>	68

Color plates: *Muraqqa'*. Album of the Indian and Persian Miniatures of the 16—18th Centuries and the Models of the Persian Calligraphy of the Same Period (see p. 63—67).

Front cover:

Fol. 17a. Portrait of a Man by Riḍā-yi 'Abbāsi, 11.8×8.2 cm.

Back cover:

Plate 1. Fol. 16a. Portrait of Timūr Khān Turkmān by Ṣādiqī beg Afshār, 19.3×11.6 cm.

Plate 2. Fol. 36a. The Darvishes Picnic in the Mountains. Probably Isfahan school, 25.5×14.5 cm.

Plate 3. Fol. 6a. The Shaykh and the Harlot by Muḥammad Yūsuf Muṣavvir, 18.2×11.3 cm.

Plate 4. Fol. 1a. Portrait of Mirzā Jalālā by 'Alī Qulī beg Jabbādār, 16.0×9.1 cm.

RUSSIAN ACADEMY OF SCIENCES
THE INSTITUTE OF ORIENTAL STUDIES
ST. PETERSBURG BRANCH



Manuscripta Orientalia

International Journal for Oriental Manuscript Research

Vol. 1 No. 3 December 1995

THESA
ST. PETERSBURG—HELSINKI

Manuscripta Orientalia

Yuri A. Petrosyan (St. Petersburg), Editor-in-Chief,
Efim A. Rezvan (St. Petersburg), Deputy Editor-in-Chief,
Edward N. Tyomkin (St. Petersburg), Editor,
Margarita I. Vorobyova-Desyatovskaya (St. Petersburg), Editor

Advisory Board

Oleg Akimushkin (St. Petersburg) — *Malachi Beit-Arié* (Jerusalem) —
Stefano Carboni (New York) — *Ronald Emmerick* (Hamburg) —
Boris Gidasov (St. Petersburg) — *Franchesca von Habsburg* (Lugano) —
Tapani Harviainen (Helsinki) — *György Kara* (Budapest) — *Anas Khalidov* (St. Petersburg) —
Evgenij Kychanov (St. Petersburg) — *Jan Mayer* (Frankfurt/M) — *Lev Menshikov* (St. Petersburg) —
Tatsuo Nishida (Kyoto) — *Giovanni Stary* (Venezia)

English Editor

Alexander Nikitin (St. Petersburg)

Technical Editor

Oleg Shakirov (St. Petersburg)

Photography

Svetlana Shevelchinskaya (St. Petersburg) and *George Skachkov* (St. Petersburg)

Copyright

© Copyright 1995 by St. Petersburg Branch
of the Institute of Oriental Studies, Russian Academy of Sciences
© Copyright 1995 by Thesa, Russia

All rights reserved. No part of this publication may be reproduced, translated, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the publisher.

Printing and distribution

Printed and distributed by Dekadi Publishing Ltd Oy, Helsinki, Finland

Subscriptions

The subscription price of Volume 1 (1995) (ca. 220 pages in 3 issues) is US\$ 132.00 for institutions and US\$ 117.00 for individuals including postage and packing.

Subscription orders are accepted for complete volumes only, orders taking effect with the first issue of any year. Claims for replacement of damaged issues or of issues lost in transit should be made within ten months after the appearance of the relevant issue and will be met if stocks permit.

Subscription orders may be made direct to the distributor: Dekadi Publishing Ltd Oy, PO. Box 976, FIN-00101 Helsinki, Finland. Tel. +358-0-638 119, Fax +358-0-638 441. Also to the publisher: 14 Dobrolyubov St., apt. 358, 197198 St. Petersburg, Russia. Tel./Fax +7(812)238-9594, E-mail bi@thesa.spb.su.

TEXTS AND MANUSCRIPTS: DESCRIPTION AND RESEARCH

L. Menshikov

A FRAGMENT OF AN UNKNOWN *LEISHU* FROM TUNHUANG

Among the Chinese manuscripts of the Tunhuang collection of S. Th. Oldenburg preserved in the Institute of Oriental Studies of the Russian Academy of Sciences there are two fragments (No. Дх-970 [1] and Дх-6116) which fit together, being a part of one scroll. The scroll contains 12 short stories about wonderful events, which happens to the following persons [2]: T'ien Chên (田真), Ts'ao Ê (曹娥), Hsün Lun (荀倫), Ling Chê (靈輒), Wei K'ô (魏顛), Wu Tzu-hsü (伍子胥), Han Hsin (韓信), Mother Chai (翟母), Yang Pao (楊寶), Sun Chung (孫種), Yang Kung (陽公), Mao Pao (毛寶).

These stories belong to the same type of fiction as the *hsiaoshuo* narratives of the Liu-ch'ao period — a very short stories about wonderful events. In the Liu-ch'ao period collections of such stories were circulated in great numbers. Many of them have been lost. It could have been possible to suggest that the manuscript considered here was a fragment of one of these collections of short novelettes, if it were not for the subtitle of one of its parts:

(報恩第二十五) “[part] 25. Requital for Favour” (line 9). Such thematic division by chapters was not characteristic of the collections of *hsiaoshuo* stories. It makes us think that this manuscript fragment belonged to some work of another genre, that is of *leishu* (類書). These were Chinese encyclopedias of a special type, consisting exclusively of texts extracted from certain works and arranged in systematic order. In this way a more or less complete selection of texts dealing with the subject in question was formed (later, in connection with the contents of the manuscript, we shall talk about the way the compilers of *leishu* were using materials collected by them).

To form a more clear idea of the genre of *leishu* let us turn to the famous encyclopedia of Chinese narrative prose “T'ai-p'ing kuang chi” [3]. The whole book is divided into several large sections titled “Men-saints”, “Women-saints”, “Taoist-Magi”, “Alchemists”, “Remarkable Buddhist monks”, “Comprehension of Buddha's teaching”, “Reward” and so on, down to animals, plants, reptiles, etc. Each section includes various stories, their contents corre-

sponding to its title. Each story is followed by a reference to its source. Other *leishu* can have different entries, references are sometimes omitted, but the same structure is maintained in all cases.

The comparison of our manuscript fragment with the scheme described above demonstrates that all the principal characteristics of *leishu* are present there: the title of the 25th section; the division of the text into two parts — a selection of stories about reverent children is followed by those about rewards for charity. Some of the stories contain references to their sources (see, for example, the story about Ts'ao Ê followed by a reference: “Extract from Tien lu”) [4].

In the Tunhuang collections in London and in Paris (collected by A. Stein and P. Pelliot [5]) and in a private collection made by the Japanese artist Nakamura Fusetsu (中村不折, 1866—1943) there are several fragments of a work written by one Chü Tao-hsing. For his work he borrowed the title of a famous collection of stories made in the 4th century by Kan Pao: “In Search of the Supernatural” [6]. In this work by Chü Tao-hsing, of which only the first *chüan* (scroll) has survived, we find all the features present in our manuscript: the name of a section: 行孝第一 “[Section] one. On Reverent Children”; some of the stories also contain references to their sources. It is possible to suggest therefore, that our manuscript presents the unknown till the present time part of the “In Search of the Supernatural” by Chü Tao-hsing. It was, apparently, a vast composition — its first *chüan* is including only the first section (it is difficult to tell, if it is a whole chapter or not), while fragment Дх-970 gives the title of the 25th section. The identification of the fragment published here with the “In Search of the Supernatural” is, of course, only a suggestion. To prove it we must wait till other fragments of the same manuscript come to light.

To confirm our attribution of manuscript Дх-970 to the *leishu* genre let us consider how the sources used by compilers are cited in the known *leishu* and in our fragment.

The first story in our manuscript tells about T'ien Chên and his brothers. It was borrowed, most probably, from Wu Chün's work [7] "Continuation of Records by Ch'i Hsieh" ("Hsü Ch'i Hsieh chi") [8]. To see how the source was

treated in our manuscript we may select parallel passages from the "Continuation of Records by Ch'i Hsieh" and from ㄉx-970. Those parts of the text which do not coincide are underlined.

"Continuation of Records by Ch'i Hsieh"

Manuscript ㄉx-970

T'ien Chên of Chingchao [and his] brothers — [there were] three persons — decided together to divide their wealth and property — all in equal shares. And in front of the house there was one tzuchingshu shrub [they] together decided to break into three parts. Next day as soon as [they wanted] to cut it, the shrub immediately got dry, as if singed by fire. Chên came, saw it and, greatly amazed, said, addressing his brothers: "The shrub by nature [is growing] from one trunk. [It heard] that we wanted to divide it, that is why it is sad. Such people [as we are] can not be compared with the tree" Therefore [they] could not hold their grief and were no longer dividing the tree. In agreement with their words the tree [again became] luxuriant and blossoming. The brothers were touched by this, joined their property and wealth and after [that] became a model family. Chên in [his] service achieved [the rank of] t'ai-chung taifu.

T'ien Chên, a man from Chingchao [and his] brothers — [there were] three of them — made a division of [their] house. And in front of the courtyard there was one shrub of chingshu. That evening when Chên with [his brothers] were dividing, [they] decided: next [morning] they would divide [it]. The branches and leaves of [the shrub] got dry and withered. The brothers saw it and, shedding tears, said: "The shrub by nature has no soul, still it is sad over this division showing us an example". The brothers after [that] were touched and [began] to live together. [T'ien Chên] was a man of the time of Wu-ti of the Han and in [his] service achieved [the rank of] t'ai-chung taifu.

The collation of these two variants shows, that our text has been considerably changed in comparison with its source, mainly in its secondary details. These details, however, are not just omitted, but are usually substituted for a brief paraphrase. For instance, a more detailed story about how the shrub withered, how the brothers saw it, and how it revived again, is substituted for a brief description in other words, but still clearly enough conveying the idea of the omitted passage.

In some cases the manuscript is supplementing the narrative of the "Continuation of Records by Ch'i Hsieh" making the information of the source more accurate. One of these supplements is the indication that T'ien Chên lived in the time of the Han emperor Wu-ti (140—87 B.C.) — a very important element of the Chinese literary tradition where literary characters were often identified with certain historical figures. Other changes are not so important. It is enough to mention that the name of T'ien Chên and the name of his native province have changed places in the manuscript (in the original source the name of the province comes first, in the manuscript it is otherwise). It was done, most probably, to attract the reader's attention to the name of the key-figure of the story. In other stories the names of the people also come first — this feature was common for *leishu*, which served as reference-books on literary plots.

The story of T'ien Chên and his brothers appears in many later *leishu*. It was included, in particular, into the famous 10th century *leishu* "T'ai-p'ing yü lan" [9]. Not citing the whole text of the story, it is enough to say that changes and reductions are present there also, though not to the same extent as in our manuscript. These are the following:

1. It is not mentioned that T'ien Chên was a native of Chingchao, so here his name also comes first.

2. Some passages are slightly abbreviated, supplemented or changed. Like in our manuscript, these changes do not affect the contents.

3. It is added, that T'ien Chên lived under the Han emperor Ch'êng-ti (32—7 B.C.). The fact that there is no mention in the "Continuation of Records by Ch'i Hsieh" of the time when T'ien Chên lived and a considerable difference in dates indicated in the 10th century *leishu* and in our manuscript, make us think that these indications are not precise. These dates could be added without any reference to historical sources, just to maintain the tradition of "connecting" the plot to some definite historical period.

In another part of the same *leishu* the story of T'ien Chên appears in a considerably more modified version [10]. Reductions and changes bring it closer to our manuscript. The presence of two different versions in the same *leishu* demonstrates that its compilers were not striving for accurate citing of their sources — their aim was just to provide the reader with a necessary reference on the plot of the story. The reader, if he wanted to get more precise information, could look for it in the original text, which was usually indicated in the *leishu*.

We see now that in the "T'ai-p'ing yü lan" the story of T'ien Chên is treated in the same way as in our manuscript.

It often happens that *leishu* are using not the whole plot, but only those parts of the narrative which are immediately connected with the subject indicated in the title of the corresponding section, omitting the rest. A characteristic example is the same T'ien Chên story cited in the 18th century *leishu* "Yüan chien lei han" in the section "Flowers" [11]. The story is even more abridged than in all other cases cited above — here the aim of the editors was to give some information about the tzuching plant, not about T'ien Chên story which happened to him. A si-

田再京北之元中三之遂不馬唯有度重一探刺樹具共示之公為
 謀議至明啟不其樹枝柔枯其妻元年見之淚淚而言曰樹下惡
 尚恨不張况我兄弟遂感而共住漢武歸人仕李太中大夫
 曾娥會誓上康之父母按江而死娥乃緣江哭之七日七夜哭聲不絕
 女之按江而死經三日其女抱屍而出家又收奔馬昂為三碑於江
 上漢時出此錄
 尚倫河內人晉時為郡守是時郡治野城倫北省舅氏某馬以盟津而
 死屍三日不得倫乃投屍於河伯經一宿其屍抱牌而出
 報恩第五
 靈輒齊人晉大夫趙道在妾下見之餓道乃候豐漿以哺之良久得蘇
 道曰子何人否曰齊人性靈名輒糗食之者豈不復能道道與與井東
 哺與之乃得深園心帝報道後任晉為守門監道每直謀靈公怒
 遣文大夫某音啞道、脚踏於下領折道謂靈公曰駭人用犬雞
 猛何為道乃走出門得破簪一、以就釘其語、輪注轉逐扶道上车
 以手成數一頭駭駕而行道得免趙道泣問之答曰賢妾下之鐵人
 把問其名后不答而返
 魏顯晉卿魏武子之子武子有寵妾武、屬其顯曰吾死後
 必嫁此妾於汝武子欲允之語顯曰必、此妾為殉及武子死顯嫁之
 從父伯之言讓秦與晉戰、顯又將夜夢見一老、為顯於第場中
 結草、執秦、軍將杜回數倒而不能去遂為顯所擒此結草
 報恩之也也出史記
 仵子曾為楚平王謀其父兄蒙水子首、巨世授美於馮飢困者於溪
 陽之水見一女子買佳從乞食女子見子、况所聞其西由胃、以餐告女子
 便與食、詭屬女請勿漏言遂辭而去可行百步見女自投、明不言
 曾仕楚王遺首將兵伐楚、伏罪秦隨子曾昂訪女家欲百金不知
 處所投金中而去該女子之母哭泣而來至水中得屍并獲金而

Fig. 1. MS. Dk-970

30
35
40
45
50

墨同第五時人

韓信淮陰世家下邵城劉其有以母怒之持郢東養食曰吾必重報之信後封為王乃曰報母報

翟母者大梁人漢高祖與項羽戰其母思之與有程安遂藏視得免難後或項羽定天下封韓信其母令陳留封立是也

楊寶私與羊季年七歲往羊陰山中見黃雀為蟻所食之妻將歸於巾中得之曰許香差之使飛去暮宿巾中相與數日化作黃衣年少持環雙報其恩謂寶曰君當貴此環累世三公後果如言三世曰於王天附漢人

孫鍾吳郡會稽種瓜為業忽有三之來就瓜鍾為摘好从其食託謂鍾曰家貧無以報恩請不子一并也將鍾上山曰吾欲得世天子欲得世世封侯鍾曰對世天子欲當為人所上三人指處可存三人曰我何命謂鍾曰君下山百步莫及願可行十步許迴看三人並作白鶴飛去後葬上有氣術天及鍾生堅字文意仕漢靈帝為心

志將軍也長文守堅生權字仲謀漢末拔江東立為吳主都揚州建業後都武昌權生亮生休為景皇帝休生皓為沖皇帝

後皓為皇帝封為歸命侯權諡曰文皇帝亮後廢合治十一年楊孚字雍伯汝南人父終葬於無名高十里公於改頭宜其狀此石嘗生玉又嘗得好婦語託曰去公移之二年往看視有玉狀此平徐公大富而女亦嫁陽公故法求之徐公笑曰卿得玉玉一可與為婚陽公於是至田取以白璧覆之遺之徐公大富為

女妻陽公此平陽早其末以漢人出故神記

毛實以西人也昔咸康中行於江道見漁人釣得一白龜實恠之腹放於江中後十年守鎮郢城與石虎身實戰敗走投江脚躡一

一龜之乃是昔時所守白龜主水守龜三視實而玉壽人

Fig. 2. MS. Dx-6116

milar case can be found in Dlx-970 manuscript: only that part of the story of Wu Tzu-hsi is cited which concerns his reward for the favour done him by the girl who drowned herself — in conformity with the title of the section.

It appears that there is nothing unusual in the way the sources are treated in the manuscript published here. It is the regular method of compiling *leishu* which, probably, is the best proof that our fragment belongs to some unknown *leishu*. All features characteristic of a *leishu* are present here.

To determine the date of the manuscript it is necessary to give its description. There is nothing in the text which may help to solve this question, so we must try to date it on the evidence of its appearance.

The manuscript presents two sheets of a scroll, their total size 72 × 27.5 cm, on the left and on the right side — traces of its being stuck to neighboring sheets. It was written on thin (0.07 mm) white paper turned slightly yellowish, with roughly ground fibers. Margins: 2 cm from the top, 1 cm from the bottom. The field of the text is equal, consequently, to 72 × 24.5 cm. There are 52 lines, each containing between 23 and 26 characters written in slightly uneven archaized *K'ai* script. On the evidence of its paper and handwriting the manuscript can be dated to the late 6th — early 7th century. It has eight yellow-brownish spots arranged symmetrically along the diagonal of the sheets, 10.5 × 2.5 cm each. Paper has partly decayed there, leaving several irregular-shaped lacunae:

1. Line 14, 10 cm from the top, lacuna 4 × 1 cm, three characters corrupted.
2. Line 17, 15 cm from the top, lacuna 5 × 2 cm, three characters missing, one corrupted.
3. Line 23, 13.5 cm from the top, two small lacunae, three characters corrupted.
4. Lines 29—30, 12 cm from the top, lacuna 6 × 3 cm, six characters corrupted.
5. Line 37, a small lacuna, one character corrupted.
6. Lines 43—44, two lacunae, 2 × 2 and 2 × 3 cm, five characters corrupted.
7. Line 49, lacuna 3 × 1 cm, two characters corrupted, one missing.

8. Line 51, two small lacunae, two characters corrupted.

The upper parts of lines 40—45 are worn out. 12 characters are corrupted in this part. There is a tear along the upper edge of line 48, where two characters are corrupted. Four or five characters are torn of the upper part of the last line and three in the middle are corrupted. There are also several small lacunae not affecting the text.

There are some other features which must be taken into account.

1. Line 3 — character 帝 omitted by the copyist is added on the right of the line. It should have been placed between characters 武 and 時.

2. Line 8 — character 求 is omitted. It is written on the right of the line, should have been placed between characters 死 and 屍.

3. Lines 12, 13, 20, 23, 24, 25 (twice), 40, 41, 43 (three times), 52 contain signs 彡 or 彣, indicating that the character preceding this sign must be repeated two times. Line 20 has this mark repeated after two characters coming one after the other which means that the combination of these two characters must be repeated.

4. Line 41 — on the left of the line, between characters 沙 and 長 there is a mark 丿. It means that the order of these characters must be changed.

By its handwriting and paper the manuscript can be dated to the 6th-7th centuries [12]. This dating makes us consider the question of the place of our fragment among the available now *leishu* [13].

It is generally believed that the first *leishu* was the book titled "Imperial Revue" 皇覽, registered in the "History of Three Kingdoms" [14] composed in 227. After that the tradition of compiling *leishu* continued uninterrupted till the 20th century. In the period in question, i. e. in the 7th century, over twenty *leishu* have been composed. None of these survived to the present time, not counting several passages from two *leishu* cited in other works [15]. Our manuscript in this way presents a fragment of one of the early *leishu* compiled not later than the 7th century. It most probably remained unknown till the discovery of Tunhuang.

Notes

1. The fragment is described in *Opisanie kitaiskikh rukopisei Dun'huanskogo fonda Instituta narodov Azii* (Description of the Chinese Manuscripts from Tunhuang of the Institute of Oriental Studies), fasc. 1 (Moscow, 1963), No. 1455.
2. In the text we use full forms of traditional Chinese characters. For the original ones see the photographs of the text, as for their explanation, see "The list of special characters".
3. 太平廣記, 李昉等編, 北京, 人民文學出版社, 1959. — Compiled in 981 by a collegium directed by Li Fang (925—996), on the order of the Sung emperor T'ai-tsung (976—998) issued in the second year of his reign under the motto *T'ai-p'ing hsing kuo*, whence from comes the title *T'ai-p'ing kuang chi*, i. e. "The Extensive Records of the T'ai-p'ing years".
4. See comments on the translation of the text.
5. P. Pelliot collection, P-2656, P-5545; A. Stein collection, S-525, S-6022.
6. 句道典, 搜神記, in: 敦煌變文集, 北京, 人民文學出版社, 1957, 865—900頁. (*A Collection of Tunhuang pienwen*) — the previously published text from the collection of Nakamura Fusetsu (羅振玉, 敦煌零拾, (no place, 1924). — Lo Chên-yü, Tunhuang finds) turned to be incomplete. The manuscript from the Nakamura collection is more extensive than the one published by Lo Chên-yü. The editors of two volumes of *A Collection of Tunhuang pienwen* published its complete version. They used, moreover, other manuscripts

from the P. Pelliot and A. Stein collections not taken into account by Lo Chên-yü. This new publication makes a good substitution for the former one.

7. Wu Chün 吳均 (469—520) — a historian, poet and writer, famous, first of all, by his historical works *The Chronicle of the Ch'i State* 齊春秋, *Records about Temples* 廟記, *Records about Twelve Districts* 十二州記, etc. His biography is included into *The History of the Liang Dynasty and History of the Southern Dynasties*, see 梁書, *chüan* 49, 四不備要, vol. 62, 上海, 中華書局, 1936. p. 238 (further "Ssu pu pei yao"); 南史, *chüan* 72: *ibid.*, vol. 69, p. 515.

8. *Continuation of Records by Ch'i Hsieh* 續齊諧記 — a collection of fantastic stories very popular in the Liu-ch'ao and T'ang periods. It is a continuation of the now lost *Records by Ch'i Hsieh* 齊諧記 — by the mid-fifth century author Tungyang Wu-i 東陽無疑. Ch'i Hsieh, whose name is present in the titles of the two books, is mentioned in "Chuang-tzu" as a connoisseur of stories about wonderful things; see: 莊子, *chüan* 1, *Ssu pu pei yao*, vol. 151, p. 3 — in the present work we use the text of the *Continuation of Records by Ch'i Hsieh* published in 秘書二十一種, vol. 20, 新安汪士漢校, 粵東文陞閣藏版, no date, p. 1b.

9. 太平御覽, *chüan* 421, 北京, 中華書局, 1960, p. 1944. Compiled on the above mentioned order of T'ai-tung (note 3). The collegium directed by Li Fang accomplished it in 983. The meaning of its title "The Imperial Survey of the Years T'ai-ping".

10. *Ibid.*, *chüan* 489, p. 2238.

11. 淵鑿類函, ch. 406, no place, 康熙四十九年, (1710), p. 2a. Compiled on the order of emperor K'ang-hsi in 1702 by a collegium directed by Chang Ying 張英 (1638—1708), first printed in 1710; on Chang Ying see *Eminent Chinese of the Ch'ing Period* (Washington, 1943), pp. 64—5.

12. It is difficult to say, if this dating can confirm our suggestion that our manuscript is a part of the *In Search of the Supernatural* by Chü Tao-hsing. None of the manuscripts of the mentioned above work is published facsimile, so we can not establish their date by handwriting and have no opportunity to compare the appearance of the two manuscripts.

13. One should take into account that in the Liu-ch'ao and T'ang periods the term *leishu* was probably not in use. It appears for the first time in *The New History of T'ang*. Before that works of this kind were attributed either to "philosophers" or to "historians", according to the traditional division of Chinese literature into four parts. See 張濂華, 類書流別, 上海, 商務印書館, 1958 (Chang Ti-hua, *A Classification of Leishu*, pp. 1—6).

14. 三國志, 魏志, ch. 2, *Ssu pu pei yao*, vol. 56, p. 39. Sometimes the Chinese tradition derives *leishu* from the Dictionary of Classics *Êr-ya*. The only feature it has in common with *leishu* is, however, the systematic order of its contents.

15. Chang Ti-hua, *A Classification of Leishu*, p. 43. Besides these *leishu* there is one more, discovered in Tunhuang and published by Lo Chên-yü 鳴沙石室古佚書, vol. 4, 京都, 1913). On the evidence of its handwriting it should be dated to the 8th—10th centuries. On its identification see Chang Ti-hua, *A Classification of Leishu*, pp. 42—3. The manner of making references to its sources is different from the one we find in our manuscript (the title of the work is preceding the passage cited from it), so it hardly could be a part of the same work to which our fragment belongs.

TRANSLATION

of Manuscripts ㄉx-970 and ㄉx-6116

[I]

(1) T'ien Chên [1], a man from Chingchao [2] [and his] brothers — [there were] three of them — made a division of [their] house. And in front of the courtyard there was one shrub of chingshu [3].

That evening when Chên with [his brothers] were dividing, [they] (2) decided: next [morning] they would divide [it]. The branches and leaves of [the shrub] got dry and withered. The brothers saw it and, shedding tears, said:

"The shrub by nature has no soul. (3) still it is sad over this division, showing us an example". The brothers after [that] were touched and [began] to live together. [T'ien Chên was] a man of the time of Wu-ti [4] of the Han and in [his] service achieved [the rank of] t'ai-chung taifu [5].

[II]

(4) Ts'ao Ê [6] from Shangyü [7] in Kueichi [8]. Her father jumped into the Chiang [9] and died. [Ts'ao] Ê was walking then along [the bank] of the Chiang, mourning for

him. For seven days and seven nights the sound of her weeping continued. (5) The daughter also jumped into the Chiang and perished. Three days passed. [The body of] his daughter embracing the corpse appeared [from the water]. The people of [their] family gave them a burial and set a [grave] stele over (6) the Chiang. [It happened] in the time of the Later Han.

Extracted from "Tien lu" [10]

[III]

(7) Hsün Lun [11] [was] a man from Honei [12]. In the time of the Chin [13] he became a district-governor. At that time the center of the district [was] in Chüanch'êng [14]. Pei, Lun's younger brother, [decided] to visit his uncle's family, mounted a horse and was drowned at the Méngchin [15] [ford] (8) and perished. For three days they could not find the body [of Pei]. Then Lun threw [into the river] an incantation letter to Ho-po [16]. One night had passed, and corpse emerged holding the letter.

(9) *Requital for Favour*

[Section] twenty-five

IV

(10) Ling Chê [17] was a man [from the realm of] Ch'i [18]. Chao Tun [19], a dignitary [from the realm of] Chin [20], saw some man near Sang [21], [who was] starving. Then Tun poured sauce from a jar to feed him. [When he] finally have had enough, (11) Tun asked: "Who are you?" The answer was: "A man from Ch'i, of the Ling family, named Chê. [Because the stocks of food] have run out [at home], I can not go back [there]". Tun gave him some grain and a bundle of (12) dried meat, so [he] could return to [his] realm, hoping in his heart to reward [Tun].

Later Ling Chê served in Chin and became an official guarding the door [of the sovereign's palace]. Every time [when there was such chance] Tun was reproaching Ling-kung [22]. Kung became angry and (13) unleashed a big hound named Fu-ao to let it tear Tun. With a kick of his foot Tun broke Ao's throat. Then Tun said, addressing Lingkung:

— A villain unleashed a hound. The hound is (14) fierce, but it is all for nothing!

After that Tun went out through the gates and wanted (to get into a cart) [23]. The socket came off and only one wheel remained. Then Chê helped Tun to get into the cart. (15) [set back] the socket with his hand, whipped up the team, and the cart started. Thus Tun managed to avoid the misfortune. Tun began to ask him in wonder. He answered:

— The hungry man near Isang (16) — is [me].

[Tun] asked him about his name and [place] of dwelling, [but he] said nothing and left.

[V]

(17) Wei K'ò [24], the son of official Wei Wu-tzu [25] [from the realm of] Chin. Wu-tzu had a favourite concubine. Not long before his death [26] Wu-tzu said, entrusting [her] to K'ò:

— After my death (18) marry this woman without fail.

Later Wu-tzu, being near his death, said again, addressing K'ò:

— Definitely, sacrifice yourself to this woman.

(19) When Wu-tzu died, K'ò married her, (19) obeying his father's words.

Then there was a war between Ch'in [27] and Chin. K'ò was appointed a commander. At night [he] saw in his dream one old man — on the battlefield [he], on behalf of K'ò. (20) was tangling grass to hold the army of Ch'in. The general of the Ch'in army, Tu-Hui, was falling down many times and could not escape. So K'ò captured [him]. This tangling of grass (21) was caused by a reward for charity [which K'ò had done by his marriage].

Extracted from "Shih chi" [28]

[VI]

(22) Wu Tzu-hsü [29]. Because P'ing-wang [30] of Ch'u had executed his father and brother [31] and announced a search for Tzu-hsü, Hsü escaped and secured himself in [the realm of] Wu [32]. On the way he starved. When he came to a river by Li (23) yang [33] he saw some woman. Hsü went [towards her] to ask for some food. The girl, seeing [that] (he was hungry, took pity on him) [34] and asked him, what had happened to him. Hsü told the whole truth. The girl (24) then gave [him] to eat. When [he] finished his meal [he] asked the girl not to let [his story become] known. He said, then bid farewell and walked further. Walking about a hundred steps he saw, that the girl committed suicide, to make it clear that [she] would never tell [anyone]. (25) Hsü served in Wu. The wang of Wu [35] sent Hsü at the head of the army to make war on Ch'u [36]. [The realm of] Ch'u collapsed, the guilty ran away. Later Tzu-hsü, on his way back, was looking for the girl's family to give them a hundred gold pieces. Not finding (26) where (her family) lived, he threw the gold into the river and departed. Later the girl's mother, weeping bitterly, came to the river. There she found her daughter's body and, taking the gold (27) went back home. Wu Tzu-hsü — a man of the time of Ching-wang of Chou [37].

[VII]

(28) Han Hsin [38], a man from Huaiyin [39]. [His] family was poor, and [he] was fishing in the Hsiap'ei [40]. One laundress desired him, took him to [her] house and fed him. [Han Hsin] said:

— I shall certainly give you a rich (29) reward for this.

Later, when Hsin received the title of wang [41], [he] invited the laundress and rewarded her with rich gifts. [He was] a man of the Early Han [dynasty].

[VIII]

(30) Mother Chai [42], a person from Taliang [43]. Kao-tsu of Han was fighting with Hsiang Yü [44]. The armies of Kao-tsu (were defeated and quickly dispersed) [45]. In Yenhsiang [46] lived Mother Chai, who concealed (Kao)-tsu, and he managed (31) to avoid misfortune. Later Kao-tsu, [when he] destroyed Hsiang Yü and established [peace] all under the Heaven, gave her (Yen) hsiang in possession to reward her for her charities. Now it is the region of Fênch'iu in the district of Ch'ênliu [47]. Mother Chai — a person of [the time of] the Early Han.

[IX]

(32) Yang Pao [48] — a man from Huayin [49] in (the district of) Hungnung [50]. [When he was] seven years old, [he] went to the Huayin mountains and saw a yellow spar-

row being eaten by ants. Then [Yang Pao] (33) took it home in a chinhsiung [51]. Fed it over ten days. Sparrow's wounds were healed and [it] took wing again. In the evening it was going to sleep in the chinhsiung. Several days later the sparrow turned into a youth clad in yellow garments, rewarded him for his charity with a pair of bracelets and said to Pao:

— I am rewarding you with these bracelets. For many generations [you and your descendants] will hold the office of sankung [52].

And it is true, that later, as had been said, in the course of three generations [he and his descendants] were suk'ungs and t'aiweis. [He was] a man [of the time] of the Later Han.

[X]

(36) Sun Chung [53] was a man from the district of Wuchün [54]. [His] family was poor, [his] occupation was pumpkin-growing. Suddenly three men appeared and began to beg [him] for a pumpkin. Chung picked a fine pumpkin. Together (37) [they] ate [it] and said to Chung:

— Being honoured with [your] food, [we] have not rewarded you for [your] charity. We wish to show you one place for burial.

With Chung they ascended a mountain and said:

— Would you like [55] [your descendants], generation after generation, to be the Sons of (38) the Heaven, or you wish them, generation after generation, to get the title of hou?

Chung said:

— Let several generations be the Sons of the Heaven and for this reason be supreme for people.

The three showed him one place, where one could be buried. Then the three (39) said:

— We command destiny, — and added, addressing Chung: — You will descend the mountain and [walk] a hundred steps without turning back.

Walking ten or more steps [he] turned back and saw how the three all turning into (40) white storks flew away. Later [Sun Chung] had been buried [there], and over [his grave] fumes were rising up to the sky.

Chung also gave birth to Sun Chien [56]. Chien, his second name Wên-t'ai, served Ling-ti of Han [57], became the Destroying (41) Captivity General [58] and the governor of [the city of] Ch'angsha.

Chien gave birth to [Sun] Ch'üan [59]. Ch'üan, his second name Chung-mou, by the end of the Han settled in Chiangtung [60] and ascended the throne as the sovereign of the [kingdom of] Wu, with its capital in [the district of] Yang (42) chou [61], in [the city of] Chienych [62]; then [he moved] the capital to Wuch'ang [63].

Ch'üan gave birth to [Sun] Liang [64]. Liang gave birth to [Sun] Hsiu [65]. Hsiu was [the emperor] Ching-huangti.

Hsiu gave birth to [Sun] Hao [66]. Hao became a deposed emperor. (43) Later the Chin [dynasty] attacked [the kingdom of] Wu, and [Sun Hao] was endowed with the title of Hou Coming back under the Authority [67].

The posthumous name of [Sun] Ch'üan was [emperor] T'ai-huangti. [Sun] Liang was overthrown. [Sun Hao's] whole reign was eighteen years [68].

[XI]

(44) Yang Kung, his second name Yung-po [69], a man from Loyang. [His] father and mother died and were buried on the mountain Wu (chung-shan) [70] rising to the height of ten li. Kung on the slope [of the mountain] set "the beverage (45) of a faithful son" [71], (to offer drink [72]) to passers by.

Three years passed. One man drank Kung's beverage, took from [73] his bosom a stone [equal to] one shêng [74] and, giving it, said to Kung:

— (46) [If you] sow this stone, jasper will grow, moreover, [you will] be rich and noble and will get a good wife.

[He] spoke and left. Kung sowed it. A year later [he] came to look at it — in the soil there was (47) something like jasper.

In Pei'ing [75] [lived] Hsü Kung [76], a very rich [man]. [He] had a daughter, not yet married. Therefore Yang Kung went to ask for her. [The man] named Hsü said laughing:

— If you get a couple (48) of jasper disks [77], it will be possible to give her in marriage [to you].

Then Yang Kung went to [his] field and, getting a couple of jasper disks, offered him. Hsü Kung was amazed and immediately (49) gave his daughter in marriage to Yang Kung. The Yang [family] from Pei'ing — are all (his progeny) [78]. [He was] a man of the Later Han [dynasty].

Extracted from "In Search of the Supernatural"

[XII]

(50) Mao Pao [79] was a man from Ch'ênliu [80]. In the years of Hsien-k'ang [81] [he] walked by the road along the Chiang and saw a fisherman who had caught a turtle. Pao took pity on it, bought [it from the fisherman] and (51) let it go into the Chiang.

Ten years later he was defending Chuch'êng [82] in the war against Shih Hu [83]. At that time Pao was defeated, ran away and jumped into the Chiang. [His] feet stepped on some (52) (stone. At the moment he) [84] looked at it — it was the white turtle saved [by him] in the former times. [Taking him across the river], the turtle submerged, three more times looked at Pao and disappeared. Mao Pao was a man of the Chin [dynasty].

Comments to the translation

1. The story of *T'ien Chên* and his brothers is borrowed from the "Continuation of Records by *Ch'i Hsieh*" by *Wu Chün* (for more information see the introductory article).
2. *Chingchao* (京兆) — an ancient district in the time of the Han dynasty. Located on the territory of the modern Shênhsi province, including the Chinese capital of that time — *Ch'angan* (modern *Hsian*), and the lands down the *Huangho* river till the region of *Huhsien*.
3. *Chingshu*, or *tzuchingshu* (紫荊樹, *ceris chinensis*) — a thorny bush which, probably in connection with this legend, is regarded as a symbol of unity in a family.
4. *Wu-ti* of the Han (140—87 B.C.) — emperor of the Han dynasty. The attribution of the life-time of *T'ien Chên* to the reign of *Wu-ti* is, most probably, arbitrary (see the introductory article).
5. *T'ai-chung taifu* (太中大夫) — a court official of high rank, standing close to the sovereign. His duty was to discuss various questions and to help the sovereign to make decisions.
6. The story of *Ts'ao Ê* is borrowed from "*Kueichi tien lu*" (see below, note 10). It also appears in "The History of the Later Han", ch. 114 ("*Ssu pu pei yao*", vol. 55, p. 1051), in "*T'ai-p'ing yü lan*" (*chüan* 31, p. 147, and *chüan* 415, p. 1916), also in some other *leishu*, always with a reference to the mentioned work.
7. *Shangyü* (上虞) — a realm in the modern province of *Chêchiang*.
8. *Kueichi* (會稽) — an ancient district which included the eastern part of the modern province of *Chiangsu* and the north-western part of the province of *Chêchiang*.
9. *Chiang* (江) — one of the names of the river *Yangtzu*.
10. "*Tien lu*" (典錄) — i. e. the "Records about *Kueichi*" (會稽典錄), the work now lost, composed at the end of the 3d or the beginning of the 4th century by *Yü Yü* (虞預, dates of life unknown). On this work see 洪煥春, 浙江地方志. 考錄. 北京, 科學出版社, 1958, 4頁.
11. No sources mentioning the name of *Hsün Lun* have been found.
12. *Honei* (河內) — an ancient district on the territory of the modern province of *Hopei*.
13. *Chin* (晉) — a Chinese dynasty reigning in 265—420.
14. *Chianch'êng* (鄆城) — an ancient city in the western part of the province of *Shantung*, not far from the modern district of *P'uyang*.
15. *Mêngchin* (盟津 or 孟津) — a ford on the *Huangho*, to the north-east of *Loyang*, the ancient capital of China (now a district in the province of *Honan*).
16. *Ho-po* (河伯) — the spirit of the *Huangho* river.
17. The story of *Ling Chê* and *Chao Tun* is borrowed from the famous historical work *Tso-chuan* (春秋左傳杜氏集解, *chüan* 10, "*Ssu pu pei yao*", vol. 5, p. 161). In "*Tso-chuan*", however, we find two independent plots, which in our manuscript are joint into one: in the original source the episode with the dog comes before the story of how *Ling Chê* saved *Chao Tun* from the wrath of *Ling-kung*, and has nothing to do with it.
18. *Ch'i* (齊) an ancient realm of China on the territory of the modern province of *Shantung*.
19. *Chao Tun* (趙盾) or *Chao Hsüan-tzu* (趙宣子), a dignitary and the chief adviser in the realm of *Chin* (see note 20) under the rulers *Hsiang-kung* (627—621 B.C.) and *Ling-kung* (see note 22).
20. *Chin* (晉) — a realm of ancient China, the central part of the modern province of *Shanhsi* and the southern part of the province of *Hopei*.
21. *Sang* (桑) or, to be more correct, *Isang* (霧桑) — hunting grounds to the south-east of the modern district of *Yütu* in the province of *Shantung*.
22. *Ling-kung* (靈公) — the ruler of the realm of *Chin*, 620—607 B.C. Enthroned by *Chao Tun* but later assassinated by him for "immoral rule".
23. There is a lacuna which partly corrupted the character 登 "to rise" which, however, is still legible, and completely destroying two or three characters following it.
24. The story of *Wei K'o* is borrowed from "*Tso-chuan*" ("*Ssu pu pei yao*", vol. 5, p. 177) with some reductions and additions, the last ones in no way changing the contents of the narrative but only explaining some obscure passages. The battle mentioned in the story took place in 597 in the eastern part of the modern province of *Shênhsi*, not far from where the *Weiho* flows into the *Huangho* river.
25. *Wei Wu-tzu* (魏武子) also known as *Wei Ch'ou* (魏擘), a dignitary and general of the realm of *Chin*, the father of *Wei K'o*.
26. There is a lacuna. Three characters are missing, only one of these — 子 (the second character of the name of *Wu-tzu*) can be restored from the context.
27. *Ch'in* (秦) — a realm of ancient China on the territory of the modern province of *Shênhsi*.
28. The reference to "*Shih-chi*" by *Ssuma Ch'ien* is, probably, a mistake, because the episode is missing in this work. It is close, however, to the corresponding place of "*Tso-chuan*".
29. The legend of *Wu Tzu-hsü*, of which the present story is a part, has two main versions: "The Biography of *Wu Tzu-hsü*" in "*Shih chi*" by *Ssuma Ch'ien* (on the flight of *Wu Tzu-hsü* see "*Ssu pu pei yao*", vol. 50, p. 761) and "The Chronicle of the Realms of *Wu* and *Yüeh*" (吳越春秋). The story by *Ssuma Ch'ien* is a narrative basing upon proved facts only. The episode where *Wu Tzu-hsü* is meeting the girl and the subsequent reward is missing there. In "The Chronicle of the Realms of *Wu* and *Yüeh*" the legend of *Wu Tzu-hsü* is rendered more fully, with some additional episodes, including the one in question. "The Chronicle of the Realms of *Wu* and *Yüeh*" obviously became the source of many later narratives and plays still circulated in China. In our manuscript we have one of the early variants of the story of *Wu Tzu-hsü*, from which only the episode with the girl is taken, slightly abridged in comparison with "The Chronicle of the Realms of *Wu* and *Yüeh*" (cf. "*Ssu pu pei yao*", vol. 115, pp. 10, 18).
30. *P'ing-wang* of *Ch'u* (楚平王) — a ruler of the realm of *Ch'u* (the territory of the present-day provinces of *Hunan*, *Hupei*, *Chianghsi*, *Chiangsu*, *Chêchiang* and the southern part of *Honan*). Reigned in 528—516 B.C.

31. The father and brother of *Wu Tzu-hsü* were executed in 522 B.C. *P'ing-wang* was arranging a marriage between his son and the daughter of the ruler of *Ch'in*, but then, captured by her beauty, himself took her as a concubine. The father of *Wu Tzu-hsü* disapproved of this unworthy deed and was thrown into prison. Fearing that his sons would try to avenge him, *P'ing-wang* ordered *Wu Tzu-hsü* and his brother to come to the court, promising to pardon their father. *Wu Tzu-hsü*'s brother came to the capital and was executed together with his father, but *Wu Tzu-hsü* escaped.

32. *Wu* (吳) — a realm in Ancient China, the territories of the province of *Chêchiang* and the southern part of *Chiangsu*.

33. *Liyang* (溧陽) — a realm on the border of the realms of *Wu* and *Ch'u*. It was located in the northern part of the district of the same name of the modern province of *Chiangsu*.

34. There are three corrupted characters in the manuscript, read as 其飢悴.

35. The *wang* of *Wu* — *Ho-lü* (閻闔), the ruler of the realm of *Wu* in 514—496 B.C. *Wu Tzu-hsü* had helped him to seize power, so in gratitude he gave him an army to attack *Ch'u* and to avenge the death of his father and brother.

36. The campaign of *Wu Tzu-hsü* against *Ch'u* took place in 506 B.C.

37. *Ching-wang* of *Chou* (景王) reigned in 544—520 B.C., which coincides only with the beginning of *Wu Tzu-hsü*'s career.

38. The story of *Han Hsin* is borrowed from "Shih-chi" by *Ssuma Ch'ien*, *chüan* 92, the biography of the *Huaiyin hou* (i. e. of *Han Hsin*), *Ssu pu pei yao*, vol. 50, pp. 927, 932. The text of "Shih-chi" is abridged and modified.

39. *Huaiyin* (淮陰) — an ancient realm in the south-eastern part of the modern district of the same name in the province of *Honan*. *Han Hsin* originated from *Huaiyin*. In 201 B.C. he received the title of the *hou* of *Huaiyin*.

40. *Hsiap'ei* (下邳) — an ancient city in the east of the present-day district of *P'eihsien* in the province of *Chiangsu*. When *Han Hsin* became the *wang* of *Ch'u*, he made *Hsiap'ei* his capital.

41. The title of the *wang* of *Ch'u* which *Han Hsin* received after the enthronement of *Liu Pang*, the founder of the Han dynasty.

42. The story of Mother *Chai* is borrowed, with some reductions, from a lost book "Legends from *Ch'ênliu*" (陳留風俗傳) by *Chüan Ch'êng* (闕稱), who lived in the 1st—2nd centuries A.D. (exact dates of his life unknown). It is cited in "T'ai-p'ing yü lan", *chüan* 202, p. 973.

43. *Taliang* (大梁) — an ancient city, not far from the present-day *K'ai-fêng* in the province of *Honan*.

44. *Hsiang Yü* (項羽) — first an ally of *Liu Pang* (*Kao-tsu* of Han) in the war against the *Ch'in* dynasty, then his rival. In the war between them (203—202 B.C.) *Hsiang Yü* was defeated and committed suicide.

45. ... were defeated and quickly dispersed ... — this place in the manuscript is corrupted, its possible reconstruction could look like 敗急散. Dynastic histories do not go into details in connection with this defeat of *Liu Pang* ("Shih-chi", *chüan* 8, "Ssu pu pei yao", vol. 49, p. 147).

46. *Yenhsiang* (延鄉) — a city near *K'ai-fêng*. Under the Han dynasty it was quickly renamed into *Fêngch'iu* (封丘). It was better known under this last name.

47. *Ch'ênliu* (陳留) — an ancient city under the present-day *K'ai-fêng*, in the Han period — the center of the district.

48. The story of *Yang Pao* first appears in the "In Search of the Supernatural" by *Kan Pao* (干寶, 搜神記, 上海, 商務印書館, 1958), *chüan* 20, pp. 151—2. The text of the manuscript is to some extent different from the available text of the "In Search of the Supernatural".

49. *Huayin* (華陰) — an ancient realm on the territory of modern *Kuanchung* in the province of *Shênhsi*.

50. *Hungnung* (弘農) — an ancient district on the bordering territories of the modern provinces of *Honan* and *Shênhsi*.

51. *Chinhsiang* (巾箱) — a box for writing implements bound in cloth.

52. *Sankung* (三公) — a general name for the three highest officials in Ancient China. In the Later Han period when, according to the "In Search of the Supernatural" and to our manuscript, *Yang Pao* lived (his name is also mentioned in "The History of the Later Han", *chüan* 46 and 80) these were: *t'aiwei*, *ssu't'u* and *ssuk'ung*, officials responsible, correspondingly, for military affairs, for accounts on population and on plots of land and for public works.

53. *Sun Chung* — the source of this story is, probably, the book titled "Good Omens" (祥瑞記) composed in the 5th—6th centuries and now lost. It is cited in "T'ai-p'ing kuang chi", ch. 389, p. 3103, where it is considerably abridged. *Sun Chung* was the ancestor of the rulers of *Wu* kingdom in the period of the Three Kingdoms.

54. The district of *Wuchün* (吳郡) approximately coincides with the territory of the ancient realm of *Wu* (see note 32).

55. *You* — in the manuscript — 吾 — "me", apparently by mistake.

56. *Sun Chien* (孫堅, 157—193), his second name *Wên-t'ai* (文臺) — governor of the district of *Wuchün*, the father of *Sun Ch'üan*, the founder of the *Wu* state.

57. *Ling-ti* (靈帝, 168—189) the last but one sovereign of the Han dynasty. He was a puppet in the hands of his favorites.

58. "The Destroying captivity general" (征虜將軍). The historical *Sun Chien* received this honorific title (its first character should be corrected to 破) for his participation in the campaign of 190, against the usurper *Tung Cho*, who actually reigned instead of *Ling-ti*, as well as his successor *Hsien-ti* (190—220).

59. *Sun Ch'üan* (孫權, 180—251, reigned 222—251) the founder of the *Wu* dynasty. His second name was *Chung-mou* (仲謀).

60. *Chiangtung* (江東) — a district in the lower course of the *Yangtzu* (*Chiang*)

61. *Yangchou* (揚州) — an ancient province including the territories of the modern provinces of *Chiangsu*, *Anhui*, *Chianghsi*, *Chêchiang*, *Fuchien*.

62. *Chienyeh* (建業) — an ancient city near modern Nanking. *Sun Ch'üan*, when he proclaimed himself emperor in 222, established there his capital.

63. *Wuch'ang* (武昌) — a city on the *Yangtzu*, now *Êch'êng* in the province of *Hupei*. There is obviously a mistake in the manuscript, because *Wuch'ang* had been *Sun Ch'üan*'s residence before *Chienyeh*.

64. *Sun Liang* (孫亮) — the ruler of *Wu*, overthrown in 258.

65. *Sun Hsiu* (孫休, reigned 258—263) — the ruler of *Wu*, known also under his posthumous title *Ching-ti*.

66. *Sun Hao* (孫皓, reigned 263—280, died in 281) — the last ruler of *Wu*. In 280 *Wu* became the possession of the *Chin* dynasty which united whole China.

67. The title "Coming back under the Authority of Hou" (歸命侯) was given to *Sun Hao* after his capitulation before the *Chin* armies and dethronement.

68. Eighteen years — this account of the years of *Sun Hao*'s reign obviously includes the year he ruled under the title of *hou*.

69. *Yang Kung* (his second name *Yung-po* 雍伯) — his story is borrowed from the "In Search of the Supernatural" by *Kan Pao*, *chüan* 11, p. 83, (with significant changes. There is, probably, a mistake in the name, because in the "In Search of the Supernatural" and in "T'ai-p'ing kuang chi" (*chüan* 292, p. 2325) he is named *Yang Yung*, his second name *Kung-po*. In the manuscript, however, the name *Yang Kung* comes through the whole story.

70. *Wuchungshan* (無終山) — a mountain in the north of the district of *Chihhsien*, the province of *Hopei*. It does not rise to the height of ten li (the more so of eighty li, as in the "In Search of the Supernatural"). The last two characters of the name of the mountain are corrupted and are not legible in the manuscript. They are restored after the text of the "In Search of the Supernatural".

71. "The beverage of a faithful son" (義漿) — in the "In Search of the Supernatural" it is explained more clearly: "There was no water on the mountain. *Kung* was collecting water and making "the beverage of a faithful son" on the slope of the mountain. All passers by were drinking it".

72. The character for "to offer drink" is corrupted and not legible. Restored after the text of the "In Search of the Supernatural".

73. The word "took from" is by mistake twice repeated in the manuscript.

74. *Shêng* (升) — a measure of capacity equal approximately to 1 liter.

75. *Peip'ing* (北平) — an ancient district in the vicinity of modern *Tsunhua*, in the province of *Hopei*.

76. *Hsü Kung* (徐公) — in the variants of the "In Search of the Supernatural" and of "T'ai-p'ing kuang chi" he is called "a man of the *Hsü* family". Here also *Kung* is probably not a name, but something like "master, lord".

77. Jasper disk (璧玉) — a circular piece of jasper with a round hole in the middle. Such natural pieces of jasper of a standard size were highly valued in Ancient China. They were probably taken for a money-account unit.

78. There are two corrupted characters in the manuscript, restored from the context as 族人.

79. The story of *Mao Pao* (died in 339) is borrowed, most probably, from the collection of stories "You ming lu" (幽明錄) by *Liu I-ch'ing* (403—444, 劉義慶). It is also cited in "T'ai-p'ing kuang chi", *chüan* 118, p. 823—4. This story is used as well in the biography of *Mao Pao*, in "The History of the Chin Dynasty", 晉書, ch. 81, "Ssu pu pei yao", vol. 58, p. 655.

80. *Ch'ênliu* — see above, note 47.

81. *Hsien-k'ang* (335—342) — the motto of the reign years of emperor *Ch'êng-ti* (326—342) of the *Chin* dynasty.

82. *Chuch'êng* (郟城) — an ancient city located near the modern district of *Huankang* in the province of *Hupei*.

83. *Shih Hu* (石虎, reigned 335—349) — a ruler of the short-lived Later *Chao* dynasty (319—349). He stormed *Chuch'êng* and took it in 339.

84. Four or five characters are torn off in the manuscript. These are restored after the texts of "T'ai-p'ing kuang chi" and "The History of Chin" as 石上與須.

**Table of the special forms of the Chinese characters
in the MSS Ⅱx-970 and Ⅱx-6116**

The following table presents forms of the Chinese characters which are differed from full standard forms or used by mistake instead of the other characters

No.	Handwriting of the MSS	Full standard forms	The combinations in which the characters used in the MSS	Line No.
1	久, 只	亦		5, 33
2	怪	(恠) 怪		15
3	經, 經	經		5, 8, 45
4	堅	聖		40, 41
5	往	往		47
6	靈	靈	靈輒, 靈公	10—13, 40
7	壺	壺	壺漿	10
8	祖	祖	高祖	30
9	直	宣	直諫	12
10	盟	(盟) 孟	盟津	7
11	權	權		41—43
12	世	世		34
13	至	至		3, 22
14	種	種		46

No.	Handwriting of the MSS	Full standard forms	The combinations in which the characters used in the MSS	Line No.
15	臺	臺	文臺	40
16	答	答	答曰	11, 15
17	召	召		32
18	指	指		38
19	晉 ^晉	晉		7, 10, 19, 43, 50, 52
20	曹	曹	曹娥	4
21	婚	婚		48
22	留 ^留	留	陳留	31, 50
23	所	所	處所	20, 23, 26, 32, 38, 52
24	件	伍	件子骨	22
25	辭	辭		24
26	衝	衝		40
27	得	得		8, 10, 12, 15, 26, 30, 37, 38, 46—48, 50
28	呼	盱		4
29	骨	骨	件子骨	22, 23, 25
30	歸	歸		5, 25, 46, 49
31	路	路		22
32	洛	洛	洛陽	44
33, I	揚	揚	揚州	41
33, II	楊	陽	楊公	44
34	與	與		30, 45, 48, 51
35	焉	焉		5
36	弟	第	弟廿五	9
37	弟	弟	兄弟	3
38	分	分	分居	1—3
39	獲	獲		26
40	哭	哭	哭聲, 哭泣	4, 26
41	笑	笑		47
42	扶	扶		14
43	虞	虞	上虞	4
44	侯	侯		38, 43
45	桑	桑	弱桑	10, 15
46	葉	葉	枝葉	2
47	來	來	來至	26, 36
48	乘	乘	乘馬	7
49	蒙	蒙	建蒙	42
50	友	犬		13

No.	Handwriting of the MSS	Full standard forms	The combinations in which the characters used in the MSS	Line No.
51	淚	淚	泣淚	2
52	輒	輒	豐輒	10, 11, 14
53	被	被		43
54	殺	殺	自殺	24
55	後	後	後漢, 於後	6, 12, 18, 19, 26, 29, 31, 35, 40, 42, 43, 49, 51
56	處	處	處所	26, 38
57	據	據		41
58	投	投		4, 5, 8, 22, 26, 51
59	沒	沒		7
60	輟	輟		15
61	走	走		51
62	是	是	於是	48, 52
63	從	從		19, 23
64	定	定		31
65	道	(道)道	趙道	10—15
66	迴	回	迴看	39
67	逃	逃	亡逃	22
68	我	我		39
69	藏	藏		30
70	地	地	京地	1
71	龜	龜		50, 52
72	此	此		18, 20, 34, 46
73	就	就		36, 45
74	示	示		37
75	貧	貧		28, 36
76	領	領	領下	13
77	願	願	及願	39
78	漂	漂	漂陽	20
79	京	京	京地	1
80	景	景	景王, 景帝	27, 42
81	征	征		40
82	恩	恩	報恩	9, 21, 31, 34
83	陰	陰	淮陰, 華陰	28, 32
84	能	能		11, 20
85	於	於	於後, 於是	5, 8, 10, 18, 19, 22, 32, 44, 48, 50, 51
86	終	終		44

THE STRUCTURE OF ISLAMIC HISTORY BOOK

(The Method of Analysis)

Among the requirements Muslim historians were expected to answer was the ability to arrange their materials so as to achieve a harmonious composition and an attractive form [1]. Among the criteria applied when estimating a work on history, the question of its structure was among those of primary importance. In this article we shall try to outline the principal forms of arranging materials employed by the medieval authors writing in Persian and Turkish and to define the place and functions of the main structural components of their works.

An accomplished historical work by a medieval Muslim author displays, as a rule, a clear structure formed by the following three components: the introduction, the main text and the conclusion. There was no established definition for the notion of "composition". Different words and combinations were used by Muslim authors to express it: *tartīb, tarkīb, naẓm, siyāq, nasaq, tansiq, tabvīb, sūrat, ilaj; tarz wa tartīb, jam' wa tartīb, sūrat-i rabṭ wa tartīb*, etc.

The introductory part usually consisted of a preface and an introduction (*dībācha, pishgustār, muqaddima*). This is the most significant part of a historical work, performing very important functions. The preface and the introduction contain information about the author, the name and the character of his work, the motives and the circumstances impelling him to write it, the time when the work was written and, often, a dedication. Sometimes the author explains there, how his materials had been collected, providing the list of his principal sources and their characteristics. Sometimes the plan of the work is given there also, along with its detailed contents (*fihrist*). This part is very important for understanding the theoretical views of the author; often the theory and the principles of historiography are explained there. There are introductory parts writ-

ten with some special aim in mind. Thus Vasif in the introduction to his book was trying to prove that after the first four righteous Khalifs there had been no sovereigns better than the Ottoman sultans [2].

Introductions to works on history are not uniform. Some are written in verse, some in prose, others in prose mixed up with verse. Usually the volume of the preface and the introduction is equal to just one page of the text, but sometimes the introductory part grows up into an independent composition which gets its own name. Thus the introduction to *Zafar-nāma* by Yazdī — on the genealogy of the Turks and the Mongols and the history of Chinghiz Khān and his descendants — became known under the title *Tārīkh-i Jahāngīr* or simply *Iftitah* ("The Beginning").

Among the questions arising in connection with medieval works on history is the sequence of their making: when the introductory parts were written — before or after the main text? There is still no definite answer supported by any convincing arguments. The solution of this problem depends, first of all, on the primary materials we select for our study. The most reliable source here can be the autographs of historical treatises. The table below contains data on seven autographs from the Manuscript Collection of the St. Petersburg Branch of the Institute of Oriental Studies.

No such elements as capital letters or division into paragraphs are present in Muslim manuscripts. Graphically the text is a single whole, with no blanks or spaces, which was dictated by a desire to use the space available as fully as possible — paper was not cheap. One of the important characteristics of the text is the equal number of lines on each folio. The course of our reasoning when making the table was basing upon these principal features of the graphic structure of Muslim manuscripts.

No.	Date	Call No.	Title	Total number of folios	Folios of preface and introduction	Number of lines		Folio and line between introduction and main text
						preface and introduction	main text	
1.	1336	C 372	<i>Majma' al-ansāb</i>	235	1b—24b	24	24	blank page between introduction and main text, 25a
2.	1516	C 491	<i>Tārīkh-i khānī</i>	350	1b—4b	12	12	4b, 4th line from below

No.	Date	Call No.	Title	Total number of folios	Folios of preface and introduction	Number of lines		Folio and line between introduction and main text
						preface and introduction	main text	
3.	1781	microfilm	<i>Durr al-akhbār</i>	132	1b—5b	29	29	5b; 13th line from top
4.	1804	C 571	<i>Firdaus al-iqbāl</i>	611	3b—16b	17	17	16b; 4th line from top
5.	1822	C 471	<i>Shāh-nāma-i 'Umar-khānī</i>	149	17b—56a	14—17	14—17	56a; 7th line from top, blank folio between preface (17b—29b) and introductory chapter
6.	1867	C 439	<i>Tārīkh-i jahān-numā</i>	843	9b—19a	29	29	19a; 4th page from top
7.	1869	C 440	<i>Tukhfāt al-tawārīkh-i khānī</i>	360	1b—15b	17	17	15b; 1st page from below

1. If introductory parts (prefaces and introductions) were written after a book had been accomplished, we could have expected to find blanks on those pages where they were joint to the main text: it is practically impossible to calculate in advance the space required for these parts and then to arrange the text on the blank pages left for this purpose strictly maintaining the number of lines per page. On the other hand, if the introductory part was written before the main text, there could be no problems with arranging the whole text, and no artificial joints between the two parts could be expected. What can the autographs of Muslim historians tell on this subject?

Let us return to the table. None of the manuscripts considered here reveal any traces of artificial joints. Pages containing the introductory part and the main text have the same number of lines; in six cases (of the seven) the main narrative starts from the same line where the introduction ends (Nos. 2—7). If we take the contents of these introductions, none of them bears any traces of a sudden interruption or deliberate prolixity — to be fitted into the space left. None of the usual techniques of condensing or rarefying of a manuscript text are used — diagonal lines, close or sprawling handwriting, etc. The free space between the introduction and the main text of *Majma' al-ansāb* is probably connected with the specific features of the manuscript itself. These are the following. The autograph has two prefaces and an introduction. Blank pages are left not only between the introduction and the main text, but also between the first and the second preface and between separate parts of the main text. The blank pages were intended, most probably, for illustrations. According to the words of the author of the manuscript, as well as from its appearance and the manner of execution, the manuscript was meant to be presented to one of the members of the Hulguid royal family reigning at that time in Iran. For some reason, however, this work has not been completed.

The presence of two blank pages (ff. 30a, 30b) between the preface and the introductory part of *Shāh-nāma-i 'Umar-khānī* is explained by the fact that the preface to this work has been left unfinished. Some free space was probably left by the author, who intended to deal with this part later. He thoroughly worked on the preface: many pages bear additions and notes on the margins, there are

corrections in the text, some part of the text on folios 25a—27b is crossed out.

2. If prefaces and introductions were written after the book had been completed, then we could have expected to find the real contents of the book in its plan drawn in the introductory part. We often find, however, when we study autographs and copies of historical treatises, that there are considerable differences between their plans and their real contents.

3. If the introductory parts of history-books were written later than the main text, then prefaces and introductions would have been missing in the books left unfinished because of their author's death or for some other reason. We find, however, that all (or nearly all) works of this kind have introductions by their authors or both an introduction and a preface. Among these are *Majma' al-ansāb*, *Firdaus al-iqbāl*, *Shāh-nāma-i 'Umar-khānī* (Nos. 1, 4, 5), also *Shajara-i turk* by Abu'l-Ghāzī and many other works.

It is evident now, that the introductory part of a work on history was written by Muslim authors in the first turn. In other words, the structural component which usually becomes the culmination of long reflections and is written by a modern author in the last turn, was serving a starting point for a Muslim historian.

In some cases doubts arise in connection with the sequence of writing of the main text and the preface. In the last chapters of *Zubdat al-tawārīkh* by Sa'dallāh A.H. 952 (A.D. 1545/1546) is twice mentioned as the date of its composition [3]. At the same time in its short (only one page) preface it is said that the title of this work is *Zubdat al-tawārīkh* and that it contains a description of events "from Adam till the time when this brief codex (*muhtasar*) was written, i. e. till year 957 of the Hegira" [4]. N. D. Mikluho-Maklay, who described the manuscript, suggested that the preface to this work had been written later than the main text. This suggestion, basing upon the assumption that the date in the preface and the date in the main text were different from the very beginning in the autograph of the work, can not be verified, because the autograph is missing. Such an inversion presents a rare exclusion. Its assumption is founded on a hypothesis, moreover that the main text of *Zubdat al-tawārīkh* mentions also A.H. 966 (A.D. 1558/1559) as the current year. N. D. Mikluho-Maklay suggested, that "this date could be just a mistake

made by the scribe when copying the work" [5]. But the same can be true in respect to 957/1550 — the presumed date of the preface.

Prefaces to the defective copy of *Al-Fuṣūl* by a 15th century anonymous author, which belongs to the collection of the Institute, and to several copies of *Kunūz al-a'zam* by 'Abd al-Rahmān Sirat are, evidently, of a later date than the main parts of these works. The reason for this is still not clear. In the words of N. D. Mikluho-Maklay the relation of the preface to the main text of *Al-Fuṣūl* "gives rise to certain doubts, and the preface to *Kunūz al-a'zam* doubtless bears traces of a later revision (by a copyist? — T. S.)" [6]. The circumstances and motives of writing a historical work could be different. There were cases when prefaces were written later than the main part. All these cases, however, must be considered as exclusive. We should try to find out why some historians were deviating from the general rule — to begin a historical narrative with the introductory *bismillāh* formula (*bī-ism Allāh ar-rahmān ar-rahīm*) and to end it with the words *tammāt al-kitāb* ("the book is completed").

Let us consider the second component of the structural triad, which was developed in historical works most thoroughly. In the Later Muslim historiography there were several ways of arranging materials of the main text. One of them was the chronological method, when the whole work was built up as a chronicle — events described according to their chronological sequence, grouped under headings like "year so-and-so" or "events of the year so-and-so". Often, when materials are sorted by dates, facts and events with no connection between them are placed side by side. The only link is the time, since all these events took place in the same year. A good illustration of this method is presented by *Tārikh-i bihān* dedicated, in the words of its author, "to the outstanding and insignificant events of history in general" [7].

Fasihi, a Timurid historian, explains the advantages of arranging materials by years: "if someone wants to find out when and where this or that event took place, he will quickly achieve it" [8]. The 14th century historian Rashid al-Dīn is, on the other hand, criticizing the chronicle pattern dominating in the Arabic historiography. In chronicles, where events are sorted by years, their connection with other circumstances of the lives of historical personalities does not become evident, so "the events do not become known as they should be, in their proper sequence and order". "Disorderly stories concerning different periods" can not be "of real use" to their readers [9].

Another way of arranging a historical work is the politico-regional. Here materials are disposed according to their geographic attribution. As an example we can take *Tārikh-i Ḥaydarī*, a vast treatise on universal history, divided into five parts (*bāb*) each of them dedicated to one region: 1) Arab world, 2) Iranian world, 3) Central and East Asia, 4) the West, 5) India. Within the chapters the chronological principle is maintained.

One more pattern is the thematic one. It was used in books on universal history as well as in monographs dedicated to some particular event or personality. *Ma'āthir al-mulūk* by Khwāndamir or *Alīhāmān-nāma-i Bukhārā* by Ibn Rūzbikhān can be taken as examples.

The dominating compositional principle in the historiography of the Later Medieval period was, however, the sorting of materials by dynasties and separate reigns. It

was applied both to universal history books and to dynastic chronicles. The chronological order was followed within each chapter.

In each case the selection of this or that compositional method was made in view of the type and subject of the work, its sources and the aims of its author. In some works we find a mixture of different structural principles and methods of composition. All these ways and methods were learnt by every new generation of historians from the works of their predecessors. The imitation of established patterns consisted in the acceptance of the plan, structure and form of the work selected. Sometimes even several different works could be used as patterns for a single compilation — one for each of its parts. According to his own words, Mirza Haydar was imitating the composition of *Zafar-nāma* by Yazdi; Hazārfann "chose the structure and form (*tarz wa tartīb*)" of *Jahān-numā* by Kātib Chālābi; historian Husayn imitated *Hasht bihish*, but in some parts of his "Wonderful Events" he "followed the same order" and arranged his materials "according to the same rules" as the author of *Tawārikh-i Āl-i 'Uthmān*; the compiler of *Tārikh-i Aḥmad-khānī* imitated, answering the request of his customer, the style and structure of *Tārikh-i Nādīrī* by Mahdi-khān Astrābādī. Versified histories by many court poets of the 15th—19th centuries were composed, according to their own words, as imitations of the 10th century poem *Shāh-nāma* by Firdausi. Compositional imitations not always resulted from a conscious act. Often compilers were "copying", not even noticing it, along with the contents of their source also its form and structure.

There are not many historical works following a "free plan" — these were produced, as a rule, by non-professional writers. Sometimes they even admit the compositional inferiority of their works [10].

The general principle of the internal organization of historical materials is their sequence in time: materials are grouped in their relation to events, and the events are arranged in chronological order. The history of mankind is divided into separate periods. In one case this division is basing upon the dynastic principle: the rise and fall of great dynasties serve the chronological frames of history. In another case the religious principle is accepted: history is divided into the pre-Islamic (*jahiliyya* — the time of "paganism", the "ignorance of the divine law") and the Islamic periods. The calendar of Islam is not uniform, even though it is connected with the events in the life of Muḥammad. In some histories it begins from "the year of the Elephant" — A.D. 570, the presumed date of Muḥammad's birth — "year two from the year of the Elephant", "year three from the year of the Elephant", etc. [11]; in some — from the year of Muḥammad's death (A.D. 632), like in *Tārikh-i Alfī*: in the majority of historical works it starts from the year of the Hegira (A.D. 622).

Events usually coincide with real space and historical time, being dated after the Hegira and the twelve-year cycle calendar (the calendar of "twelve animals"). The Islamic era is used more often; in several cases both systems are applied, which sometimes leads to contradictions in dates. Thus, according to Naṭanzī who wrote in Fars, Al-guy "died in 664/1265—1266, corresponding to the year of the Pig" [12]. In reality A.H. 664 corresponds to the year of the Panther. Mahmūd b. Valī, a Central Asian historian of the 17th century, is mentioning the year of the Mouse (A.D. 1468) as the year of the Shibanid ruler Abu'l-Khayr-

khān's death. At the same time he is dating this event to A.H. 874, corresponding to A.D. 1469/1470 [13]. In one Turkish version of *Tārīkh-i Rashīdī* it is mentioned, that this translation has been done in East Turkestan "in the year 1160, corresponding to the year of the Sheep" [14]. In reality not 1160 but 1164 (A.D. 1751) corresponds to the year of the Sheep.

In similar cases of chronological discrepancies present in works written in Turkestan "one should presume a mistake in the Muslim date, because the natives of the land were more familiar with the animal-cycle calendar" [15]. This subtle observation made by V. V. Barthold was basing upon a long and thorough study of all the Muslim sources available at that time. When we apply it in practice, however, we must take into account one fact not considered by V. V. Barthold: in Turkestan and its neighborhood the twelve-year cycle calendar was not uniform. It is testified by several historians. Mahmud b. Vali is mentioning, though not with full confidence, that there was some difference between the calendar "day and week" in Kashgar and those in Maverannahr, Balkh and other regions. He is referring to "a rumour" [16].

The evidence recorded in *Tārīkh-i amniya* by a Turkestan historian Mulla Mūsā is more definite. In his work it is said that in the reign of Sa'īd-khān (1514—1533) a correction was introduced into the twelve-year cycle calendar used in "the Seven Cities". The result was that "in these cities the reckoning of years is by four years ahead of the reckoning of other cities. The (Hegira) dates there are, however, the same as elsewhere" [17].

Evidently, it is not enough to refer to the available synchronistic tables when translating dates of one calendar into another system of reckoning. The examples cited above show, that local "corrections" must be taken into account to explain discrepancies between the Hegira dates and the years of the twelve-year cycle calendar.

One should notice certain particularity in translating dates of the Muslim calendar into the European system of reckoning. Sometimes, when dates are translated to an approximation of the day of the month and of the week, there occur discrepancies within one day. The reason for these chronological problems is not the inaccuracy of Muslim authors or the inadequacy of the method of calculation but the difference between the European and the Medieval Muslim way of reckoning days. According to the European tradition every new day begins at 12 o'clock p.m. ending after 24 hours; in the Muslim tradition a new day begins immediately after sunset and continues till the next one. In that way every day of the Muslim calendar coincides with two days of the European calendar.

The Hegira dates are usually written in Arabic. Chronological dates written in numerals rarely occur in Muslim manuscripts. Sometimes dates are recorded as chronograms composed in a way that their letters summed up according to the *abjad* (i. e. to their numeric values) make the required Hegira date. There are some curious cases. For example, in one of the copies of the Turkish translation of *Tārīkh-i Rashīdī* the name of the Khazakhs is written sometimes as qazāq, sometimes as qadāq. It continues till page 229. This page contains the passage: "... one scholar composed a chronogram on this event: Āshti//qadāq. According to the *abjad* the date required (A.H. 919 of the Persian original) could be obtained only from the combination āshti//qazāq. For this reason the word qadāq is

crossed out and qazāq inscribed above it. After that only the word qazāq is used for "Khazakh" in the rest of the text [18].

Sometimes, to improve the general structure of the text, its author was changing the sequence of his narrative deviating from the usual chronological order and confusing it. Starting to describe one thing he is freely changing the subject; often one and the same theme is repeated at different places, the narrative is interrupted by recollections and additions often ending in words like "let us now return to our story". Additions-digressions from the main subject are sometimes so vast, that the author is obliged to introduce after them a special chapter titled "Back to the principal story" surveying in brief the contents of the chapter interrupted by this insertion. Sometimes "inserted works" — usually small compositions by other authors — are introduced into the main text. There are, for example, three insertions of this kind in *Tārīkh-i Rashīdī*. Their contents is in no way connected with the main text. On the other hand, they can not be treated as something alien to the structure of *Tārīkh-i Rashīdī*. Two of these compositions belong to the spiritual guide of Mirza Haydar — the author of *Tārīkh-i Rashīdī*. These insertions present a compositional method answering with the author's intention to give his readers some more hints on his personality and his biography.

In some historical works there are words addressed by their authors to the reader, urging him to supplement their works by a description of contemporary events. In the world-history *Ibrat al-nāzīrin* by Māsīhi the narrative is brought down to the time of Timūr's death. At the end of the book it is said, that if the author has more years to live, he will write also the history of the events following, if not — let anyone, who is able and willing, add it to his work [19]. Sometimes this appeal was getting a response from "those able and willing": among historical works there are books with *dhayl* (addition, supplement) written by a different author. Often the composition of a *dhayl* follows, in its turn, the familiar three-fold scheme: a preface, the main text (supplement) and a conclusion.

There is one more structural peculiarity which occurs, in particular, in Central Asian histories. It is known that among the Central Asian dynasties of the 15th—19th centuries only the Shibanids and the Ashtarkhanids, the khans of the Khazakhs and of Khiva, could claim to be the descendants of the family of Chinghiz. Timūr and the Timurids (1370—1506), the Mangyt dynasty of Bukhara (1785—1868) and the rulers of Kokand of the Ming dynasty (1709—1876) were not connected to the Chinghizids. Meanwhile the exclusive right to ascend the khan's throne belonged only to the Chinghizids, the descendants of Chinghiz through the male line. Any emir who managed to seize power and to found a new (non-Chinghizid) dynasty was striving either to marry a khan's daughter (like amir Timūr) and thus to obtain the honourable title of *ghurghan* (khan's son-in-law) [20] — in this case court historiographers could compose a fantastic genealogy for him; or he could add additional links to his genealogy, thus connecting his family with one of the dynasties previously reigning in Central Asia (like in the case of the Ming dynasty). It was not only a tradition. The necessity of such substantiation was recognized in official circles. Genealogy was an important factor in the state and political life of that time. The tradition to choose khans only from the

members of the Chinghizid family continued till the middle of the 19th century.

This tradition is reflected in several dynastic histories. It affected, first of all, the introductory parts of these chronicles. Many dynastic histories written under Timur or the Timurids, Shibanids, Ashtarkhanids, the khans of Khiva have introductory chapters containing a brief survey of political events in Central Asia preceding the reign of the dynasty or the ruler — the main subject of the narrative. This survey usually begins with the story of Chinghiz Khān's conquests or contains the pedigree of the ruler to whom the work is dedicated, its principal aim — to prove this ruler's connection with the Chinghizid family. Historical works by scholars from Ferghana start with the genealogy of the khans of Kokand, of the Ming dynasty, tracing it to Timūr through Babur, and in this connection relating the events of Babur's time (1483—1530). In the history by Qadir 'Ali-beg (17th cent.), who came from the Khazakh tribe of Jālair, much attention is given to the genealogy of Khazakh khans and sultans who, according to the tradition, were descendants of Tūghā-Timūr, the grandson of Chinghiz Khān.

The material of historical books usually presents a text written in prose or versified. Few authors were introducing tables or various graphic figures as "visual aids". The major part of the texts of *Mu'izz al-ansāb* and *'Umdat al-tawārikh* represents rows of circles, large and small, with names written within them, and with historical explanations in prose arranged nearby. Several folios are covered with circles containing tribal names, other pages show tamgas (heraldic devices) of different tribes, etc. [21] The majority of works on history contain verse, both by their authors and by other poets, inserted into the prosaic text. Usually these versified insertions serve only as literary decorations, though some of them contain valuable information.

Any historical narrative is subject to the inner logic of events, so its text has, as a rule, a multistage structure. It can be divided into volumes (*jild*, *mujallad*, *kitāb*) or parts (*qism*, *daftar*, *rūkn*, etc.), chapters (*bāb*, *fasl*, *rauza*, *dāstān*, *tābaqa*, *hikāyāt*, etc.) or sections (*maqāla*, *tā'ifa*, *jumla*, *chamana*, etc.) which, in their turn, may be split into smaller fractions beginning with words *dhikr*, *bayān*, *guftār*, *qissa*, etc. The main principle of division of the main text is, however, by chapters. There is no strict order regulating the use of all these terms. Often definitions used for large structural units by one author, are employed by another to indicate smaller units, or vice-versa.

The structural function of sections can be different. The medieval Moslem historians were sometimes explaining the reason for this division of their works into volumes and parts and into a certain number of chapters. At the end of the St. Petersburg copy of *Tārikh-i jahān-numā* it is said that its author, considering the formidable size of his work, "decided to divide it into two volumes (*daftar*) for the convenience of those reading and understanding (the

book)" [22]. In the words of Mirza Haydar, his intention was to write the history of the Mongols. After a long consideration over the composition of his work, he decided to divide it into two *daftar*s. The first one was to contain everything obtained from written sources and from recollections, the second — to describe the events of which the author himself was the evidence. According to this plan each *daftar* presented a compositional whole, including, besides the main text, a preface and a conclusion. Each part was subject to the general idea of the work. The historian claims, that by arranging his materials in that way he could most efficiently realize his intention [23].

Abu'l-Ghāzi in the preface to his work tells, that he called it *Shajara-turk* and distributed his materials among nine chapters (*bāb*). "These nine chapters made a wonderful conformity, because, as the wise men say: 'the degree of all being does not exceed nine. Nine is a limit to everything'" [24].

We should mention, that the initiative of dividing a book into volumes and smaller units not always belonged to its author. The will of his customer was a decisive factor in the choice of the structural form of his work.

Historical books usually have an epilogue (*guftār dār iḥtītām*), a conclusion (*hātima*, *iḥtītām*) or a supplement (*tatimma*, *dhayf*). These final parts of historical works are far from being uniform. Some contain a geographic supplement, some — biographies of the celebrated, from the author's point of view, people of his time or of the author himself, and some — just entertaining stories. There is no case, however, when an author repeats the contents or gives a summary of the main text of his work. The word "conclusion" as we understand it — the summary of author's results and principal ideas — can not be applied to the works of medieval Muslim historians.

Materials on the structure of a Muslim historical treatise considered here in general terms can be, as a whole, interpreted in the following way. A high level of unification, even of standardization of various kinds of historical books, conservatism and almost total absence of structural development are characteristic of the medieval Muslim historiography. Serious changes in the established patterns of thinking and in the way of presenting materials are comparatively rare. The process of structural development of historical sources was connected, first of all, with a transition from "universal histories" to regional and dynastic chronicles. This, doubtless, affected the structure of a book. Often it led to local tendentiousness, making the outlook of a historian more provincial. It is natural that the methods of approach to historical materials developed in the course of making of Muslim historiography turned to be very stable and uniform for the whole Islamic world. These circumstances, in our opinion, allow us to apply effectively the method of comparative historical analysis (including typological comparison) to apparently different historical works written at different periods, in different regions and in different languages.

Notes

1. T. I. Sultanov, "Vzgliady pozdnesrednevekovykh musul'manskikh avtorov na istoricheskuiu nauku" ("The views of the late Medieval period authors on science of history"), *Narody Azii i Afriki*, 1 (1988), p. 56.
2. V. V. Bartold, *Sochineniia* (Works), vi (Moscow, 1966), p. 315, footnote 73.
3. Sa'dallah ibn 'Abdallah, *Zubdat al-tawārikh*. Manuscript of the St. Petersburg branch of the Institute of Oriental Studies, call number C 505, ff. 330b, 340a.

4. *Ibid.*, f. 2a.
5. N. D. Mikluho-Maklaĭ, *Opisanie persidskikh i Tadzhikskikh rukopiseĭ Instituta vostokovedeniia AN SSSR*. Fasc. 3. *Istoricheskie sochineniia* (The Description of Persian and Tadjik Manuscripts of the Institute of Oriental Studies, USSR Academy of Sciences. Fasc. 3. Historical Writings) (Moscow, 1975), No. 315.
6. *Ibid.*, No. 420.
7. *Tāriḫ-i bihān*. Manuscript of the St. Petersburg branch of the Institute of Oriental Studies, call number C 458, f. 131b.
8. Faṣiḥ Aḥmad b. Jalāl al-Dīn Khavāfi, *Mujmal-i Faṣiḥi*, i (Mashhad, 1341/1963), p. 8.
9. Rashid ad-Din, *Sbornik Letopiseĭ* (The Collection of Chronicles), i, 2 (Moscow—Leningrad, 1952), p. 73.
10. ‘Awad Muḥammad, *Tāriḫ-i jahān-numā*. Manuscript of the St. Petersburg branch of the Institute of Oriental Studies, call number C 439, ff. 834b, 835a.
11. *Mujmal-i Faṣiḥi*, i, pp. 41—64.
12. *Extraits du Muntakhab al-tavarikh-i Mu’ini* (*Anonyme d’Iskandar*), publiés par Jean Aubin (Teheran, 1957), p. 104.
13. Mahmūd b. Vali, *Baḥr al-asrār*, vi, part 3, f. 160b (microfilm FV 82 of the St. Petersburg branch of the Institute of Oriental Studies).
14. *Tāriḫ-i Rashidī tarjamasī*. Manuscript of the St. Petersburg branch of the Institute of Oriental Studies, call number C 570, f. 240b.
15. V. V. Bartold, *Sochineniia* (Works), vii (Moscow, 1971), p. 299.
16. Mahmud ibn Vali, *More taĭn otositel’no doblesteĭ blagorodnykh (geografiia)* (The Sea of Mysteries Concerning Valours of the Noble. Geography). Preface, translation, notes, index by B. A. Ahmedov (Tashkent, 1977), p. 72.
17. *Materialy po istorii Kazakhskikh khanstv XV—XVIII vekov (Izvlecheniia iz persidskikh i tiurkskikh sochineniĭ)* (Materials on the History of the Kazakh Khanates of the 15—18th Centuries. Extracts from the Persian and Turkic Writings) (Alma-Ata, 1969), pp. 482—3.
18. *Tāriḫ-i Rashidī tarjamasī*. Manuscript of the St. Petersburg branch of the Institute of Oriental Studies, call number C 569, p. 229 (European pagination).
19. V. V. Bartold, *Sochineniia* (Works), viii (Moscow, 1973), p. 406.
20. It is noteworthy that Mahmūd b. Vali is applying the word *Ghurghaniyan* to Timūr’s dynasty (Mahmūd b. Vali, *op. cit.*, p. 7).
21. Mir Rabi’, ‘*Umdat al-tawāriḫ*. Manuscript of the St. Petersburg branch of the Institute of Oriental Studies, call number B 1876, ff. 45b, 46a ff.; *Mu’izz al-ansāb*. Manuscript of the Bibliothèque Nationale de Paris, call number P. 67, ff. 6, 12, 24 ff.
22. ‘Awad Muḥammad, *op. cit.*, f. 835a.
23. T. I. Sultanov, “‘*Tāriḫ-i Rashidī*” Mirzy Haĭdara Duglata (literaturnaia istoriia pamiatnika) (“‘*Tāriḫ-i Rashidī*” by Mirza Ḥaydar Dughlāt. Literary History of the Text”), *Pis’mennye pamiatniki Vostoka. Istoriko-filologicheskie issledovaniia* (Moscow, 1982), pp. 121—2.
24. *Histoire des Mogols et des Tatars par Aboul-Ghazi Bēhadour Khan*, publiée, traduite et annotée par le Baron Desmaisons, i (Texte) (St. Petersburg, 1871), p. 4.

TO THE HISTORY OF ORIENTAL TEXTOLOGY

K. B. Kepping

THE OFFICIAL NAME OF THE TANGUT EMPIRE AS REFLECTED IN THE NATIVE TANGUT TEXTS

In the last decade academic interest in Tangut civilization has rapidly grown and recently (August 1995) even a special conference on Tangut problems was held in Yinchuan, China [1]. This is explained by the fact that many scholars throughout the world became aware both of the unique and extremely high level of Tangut culture, as well as its significance in the history of Central Asia. However, not long ago Tangut studies were regarded as one of the backwaters of Oriental studies, while the Tangut state was supposed to be an Empire in quotation marks, which did not play any significant part in history (Clauson, 1964: 54). Such a change in the approach to the Tangut culture is due to the systematic work of an international team of scholars whose mutual efforts have created a new branch of Oriental studies, namely — tangutology. In this connection first of all the great Russian scholar N. A. Nevsky (1892—1937) is to be remembered, since precisely his works have laid the foundation of modern tangutology.

However, it must be kept in mind that now almost everything we know about Tangut is based on Chinese source material (mainly on Chinese dynastic histories). At the same time hundreds of Tangut texts are kept in different parts of the world, the largest collection — the Kozlov collection in St. Petersburg, Russia — having more than ten thousand items. Tangut collections usually consist of texts translated from other languages. For tangutology these texts are not as important as the native Tangut texts are. By native Tangut texts I mean the texts not translated from other languages, but primarily written in Tangut language (Tangut code, poetry, proverbs, official documents, etc.). These texts are kept only in the Kozlov collection in St. Petersburg. However, and specialists know it quite well, up to now some of the native Tangut texts are still beyond comprehension due to their extremely complicated content.

My assertion that native Tangut texts are of exceptional significance is corroborated by the fact that only in these texts one can find the official name of the Tangut kingdom — (1) [2] *phon mbin lhie tha* "The Great Kingdom of the White and Lofty". This name is not to be found in Chinese dynastic histories where the Tangut Empire is referred to as *Xia Guo* "the Xia State" (*Song shi*) or *Xi Xia* "the Western Xia [State]" (*Liao shi* and *Jin shi*). The character *xia* "summer" in the name of the Tangut Empire goes back to the ancient Chinese dynasty *Da Xia* (III—II millennium B.C.). Later in the 5th century A.D. on the territory of the *Da Xia* was situated the state of Xiong-nu(?)

prince Ho-lien Po-po and his state was also named *Xia*. Thus, it is quite clear that *Xia* is the name which Chinese (= the foreigners) used to designate the Tangut Empire, whereas (1) *phon mbin lhie tha* "The Great Kingdom of the White and Lofty" obviously represents the name Tangut made themselves. It goes without saying that without an adequate understanding of the meaning concealed by the Tangut in the official name of their state one will get a wrong idea about the whole Tangut culture.

In Tangut texts one can find two names of the Tangut Empire: the one I have already mentioned (1) *phon mbin lhie tha* which character by character denotes "white", "high, lofty", "kingdom", "great" (henceforth: Name 1) and (2) *phon mbin tha ndziwe lhie* which character by character stands for "white", "high, lofty", "great", "summer", "kingdom" (Name 2). The both Names are almost similar. Each contains the noun (3) *lhie* "kingdom, state, realm" and the three adjectives (4) *phon* "white", (5) *mbin* "high, lofty" and (6) *tha* "great". A diagnostic sign of the class of adjectives in Tangut language is their position with respect to the noun they modify. As a rule, in Tangut the adjective follows the noun it modifies. The difference between Names 1 and 2 lies, first of all, in the "superfluous" character (7) *ndziwe* "summer" (= *xia*) and, secondly, in the position of the adjective (6) *tha* "great", which in Name 1 modifies the noun (3) *lhie* "kingdom", but in Name 2 modifies the noun (7) *ndziwe* "summer" (*xia*): (8) *ndziwe tha* "The Great Kingdom" in Name 1 vs. (9) *tha ndziwe lhie* "The Kingdom of the Great Xia" in Name 2. Name 1 and Name 2 have their own domain of usage. In the collection of Tangut and Chinese texts published by Shi (1988: 231—330), which comprises prefaces and epilogues to Tangut translations of Buddhist and other works, only Name 1 occurs in Tangut texts prior to 1227, *i. e.* at the time the Tangut state itself actually existed. What is more important, only Name 1 is used on the stelae erected near the tombs of the Tangut emperors. It is Name 1 which stands at the beginning of the Tangut translation of the Buddhist Canon (*fig. 1*). And the famous Liangzhou stela begins with the Name 1 (Nishida 1964: 161). Name 2 does not occur in such texts, but occurs exclusively in Chinese texts written in the Tangut state. Name 2 began to appear in Tangut texts only after the fall of the Tangut Empire. This is only natural in view of the fact that the collocation (10) *tha ndziwe* "Great Xia", as we already know, constitutes the Chinese name for the Tangut Empire. Since

the Tangut Empire encompassed the territory once occupied by the Da Xia dynasty, the Chinese used the historical name Xia to describe the Tangut State, whereas this name evidently meant nothing to the Tangut themselves. It should be noted that in Tangut poetry the word (7) *ndziwe* “summer” is used only as a designation for the season itself. Also quite typical is the purely Chinese word order in the combination (10) *tha ndziwe Da Xia* “Great Xia”, where the adjective precedes the noun it modifies. It is to be stressed that Name 1 is aesthetic and well-balanced: (1) *phon mbin lhie tha* “The Great Kingdom of the White and Lofty” (mind the Tangut word order in the collocation (8) *lhie tha* “Great Kingdom” — lit. “Kingdom Great”), whereas I consider Name 2 to be awkward and un-Tangut-like in that it combines two disparate conceptions of the state, i. e. both the Tangut and Chinese notions: (2) *phon mbin tha ndziwe lhie* “The Kingdom of the Great Xia of the White and Lofty”. This is why I believe that only Name 1. (1) *phon mbin lhie tha* “The Great Kingdom of the White and Lofty”, is the native and traditional name the Tangut made for their country, whilst Name 2 represents a Sino-Tangut “hybrid”. The usage of the Name 1 on the stelae near the tombs of the Tangut emperors shows that it represents the official name of the Tangut State which contrasts with the unofficial name of the state — (11) *mi lhie* “The Kingdom of Mi (= Tangut)” used in the text of Tangut Code. The name of the Tangut State on the title-page of the Tangut Code would be very significant, but, alas, the title-page is lost (Kychanov 1987b: 245).

At the beginning of the 1930's scholars became acquainted with the Tangut original names for their country. The difference between the two names for the Tangut state has not been a source of controversy among scholars, all of whom agree that (7) *ndziwe* “summer” points to the Xia dynasty. As for the shared component in the two names, viz. the collocation (12) *phon mbin* “white and lofty”, all attempts to arrive at consensus have failed, and the issue has given rise to a polemic amongst scholars. The debate on the meaning of the collocation (12) *phon mbin* “white and lofty” was opened by Wang (1932), who asserted that in the name of the Tangut country the character (5) *mbin* was used phonetically with the value [mi]. Wang made the supposition that the Tangut word for “man” must have been pronounced something like [mi], as in the Tibetan word *mi* “man”. According to his view, the collocation in question is purported to signify “White Mi”, which was supposed to have meant “White People”, leading Wang (1932: 77–88) to the conclusion that Name 1 should be translated as “The Great Kingdom of the White Mi”. Accordingly, Name 2 would have to be translated as “The Kingdom of the Great Xia of the White Mi”. Wang's contention is all the more remarkable in that it raises the question why the Tangut did not in this context use the character (13) *mie* “man”, (14) *mi* “Tangut” or (15) *mi*, the first component of (16) *mi niau* “Tangut” (from Tibetan *mi-nag* “Tangut”), all of which sounded similar to Tibetan *mi* “man”, the latter two even specifically denoting “Tangut”. Nevsky took issue with Wang's interpretation in his famous “O nazvanii tangutskogo gosudarstva” (“On the Name of the Tangut State”) (1960: 1, 33–51) in which he meticulously analyses Wang's theory and advances a number of arguments which disprove it. In particular, Nevsky rightfully points out that if the combination (12) *phon mbin* were to signify “white people”, the adjective (4) *phon*

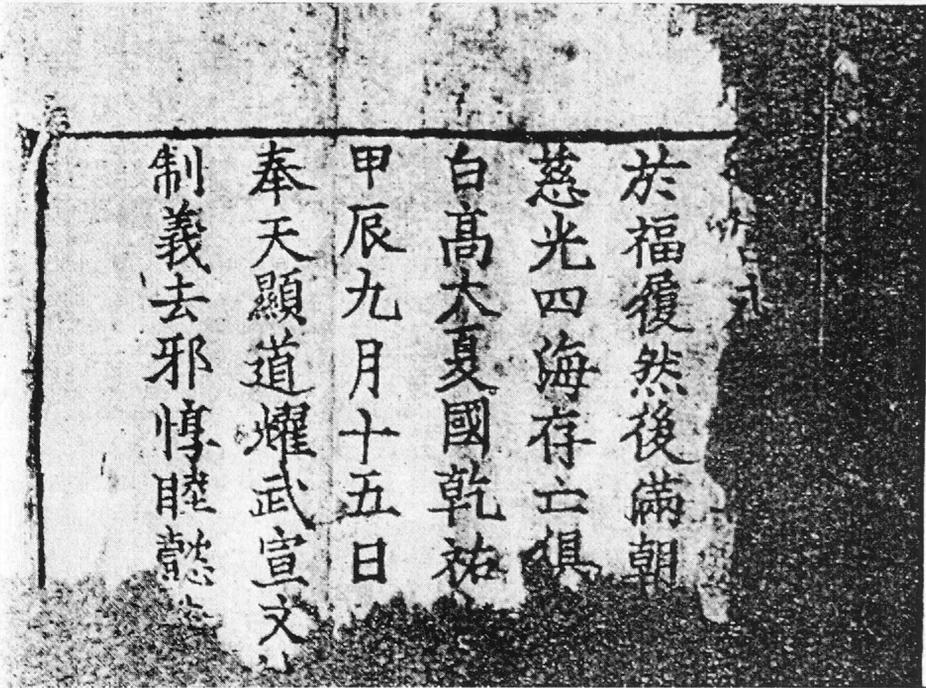
“white” would have to follow the “noun”, which it does not. Nevsky made use of the ode “Hymn to the Sacred Ancestors of the Tangut”, a previously unstudied Tangut native text later published by Kychanov (1970: 217–31), and the Tangut encyclopaedia “The Sea of Meanings Established by the Saints”. The first three lines of the “Hymn to the Sacred Ancestors of the Tangut”, in Nevsky's opinion, define the location of the Tangut State. These lines are quoted below with Nevsky's translation [1] and my translation [2]:

(17) <i>u</i>	<i>nia</i>	<i>ldu</i>	<i>we</i>	<i>tshon</i>	<i>zie</i>	<i>mbiu</i>
head	black	stone	town	desert	river	beside
<i>ni</i>	<i>nin</i>	<i>via</i>	<i>lhio</i>	<i>ma</i>	<i>phon</i>	<i>mbin</i>
face	red	father	tomb	river	white	high
<i>mi niau</i>	<i>ndzio</i>	<i>'in</i>	<i>lhie</i>	<i>tha</i>	<i>viei</i>	
Tangut	tall	possessive	state	here	is	

- [1] The stone city of the black-headed ones on the banks of the desert waters.
The paternal burial mounds of the red-faced ones along the upper reaches of the White River.
Here is to be found the land of the tall *mi-niau* (viz. Tangut).
- [2] The stone city of the black-headed ones on the banks of the uninhabited river.
The paternal burial mounds of the red-faced ones [at the foot of the mountain] the “White and Lofty Mother”.
Here is to be found the land of the tall *mi-niau* (viz. Tangut).

On the assumption that the first two lines of the ode are precisely parallel in construction, Nevsky juxtaposes the three-ideogram collocation in line 1 (18) *tshon zie mbiu* “desert river beside” to the three-ideogram collocation in line 2 (19) *ma phon mbin* “river white high”.

Nevsky did not actually call the words (20) *mbiu* and (5) *mbin* postpositions, but it is evident from his analysis that he perceived them as such. Nevsky maintained that the word (5) *mbin* “high, lofty” in line 2 meant “upper reaches”, corresponding to Chinese *shang* (later this proved to be incorrect, see below), and that the three-ideogram phrase (19) *ma phon mbin* is to be translated properly as “upper reaches of the White River”, whereby Nevsky isolates the phrase (19) *ma phon mbin* in the meaning “upper reaches of the White River”. Nevsky identifies the “White River” with the Bai Shui in the Minshan mountains in modern Sichuan province. He arrives at this conclusion for two reasons: 1) The Bai shui is believed to have been the habitat since the Han dynasty of the Qiang tribes, the ancestors of the Tangut. Subsequently, according to Nevsky, when later the Tangut had established their own realm they named it after their native homeland where lay the ancestral graves of the Tangut. In this way, their new state came to be called “The Great Kingdom of the Upper Reaches of the White [River]”; 2) In the “Hymn to the Sacred Ancestors of the Tangut”, the borders of the Tangut state are indicated, with the Minshan mountains constituting the southern border. Nevsky (1960: 1, 48) mentions that the prominent Chinese scholar Luo Fucheng also interpreted (19) *ma phon mbin* as *Bai Shang* “upper reaches of the White [River]”. According to Luo, this name could



於福履然後滿朝
慈光四海存亡俱
白高大夏國乾祐
甲辰九月十五日
奉天顯道燿武宣文
制義去邪惇睦懿

Fig. 2

be taken either as indicating that the Tangut lived to the west of the loop of the Yellow River, in other words, along its upper reaches where its waters are not yet yellow, or, alternatively, that the Tangut lived along the upper reaches of the White River, *i. e.* the Bai Shui.

In the 1960s E. I. Kychanov again raised the question of the interpreting of the name of the Tangut State. E. I. Kychanov has found among Chinese texts published in the Tangut state a Chinese translation of the name of the Tangut Empire — *Bai Gao Da Xia Guo* (fig. 2). From that time on we have at our disposal an authentic (not a later reconstruction) Chinese name for the Tangut State used in the times when this empire was still alive. And, as we see, in this Chinese name the Tangut character (5) *mbin* is rendered into Chinese as *gao* “high, lofty” (not *shang*!). Supporting Nevsky in his criticism of Wang, E. I. Kychanov nevertheless disagreed with what he maintains is one of Nevsky’s main premises, *viz.* that the Tangut State’s southern borders extended as far south as Sichuan. The fact that the ancestors of the Tangut lived in Sichuan does not necessarily imply that the Tangut would centuries later link the name of their realm to that of a river in their ancestral homeland. E. I. Kychanov (1968: 53–54) casts doubt on the interpretation of (12) *phon mbin* as “upper reaches of the White [River]”, emphasizing the fact that the word “river” itself is left out of the text entirely. E. I. Kychanov does not provide his own interpretation for the name of the Tangut State. Rather, based on a translation of the name for the Tangut State by Chinese contemporaries, *Bai Gao Da Xia Guo* “White High Great Xia State”, E. I. Kychanov (1968: 55) proposes that the Name 2 for the Tangut realm be tentatively translated as the “White and High Kingdom of the Great Summer”. Later, however, E. I. Kychanov without explanation translates the name of the Tangut kingdom as the “White High Great Hsia” (1978:25), the “White, High and Great Empire” (1984: 386), and as “The Great White and High Kingdom” (1987a: 134). But in E. I. Kychanov’s translation of the Name 1 there is a grammatical mistake. If his translations were to be correct, in which all three adjectives shared the same syntactic status as attributes to the noun (3) *thie* “kingdom”, then all three adjectives would have to occupy the same position with respect to the noun and would, more specifically, have to come after the noun. Instances of three adjectives in succession following a noun are attested in Tangut (Kepping 1985: 384, ex. 951). Of the three adjectives in the name for the Tangut realm, however, only one — (6) *tha* “great” — follows the noun, whilst the other two precede it. In Tangut grammar the different positions of these adjectives with respect to the noun unequivocally indicate a difference in syntactic function. The translations “White High Great Hsia”, “The White, High and Great Empire”, “The Great White and High Kingdom” for (1) *phon mbin thie tha* are therefore incorrect. In the 1990s E. I. Kychanov gave another translation for the collocation (12) *phon mbin*, *viz.* “white heights” (1993: 49, 53; 1994: 94). However, this translation also proves to be grammatically incorrect, since, first, (5) *mbin* “high, lofty” is not a noun, it is an adjective, and, secondly, if (5) *mbin* were a noun, then the adjective (4) *phon* would have to follow the “noun”, which it does not.

Most modern Chinese scholars are of the opinion that the combination (12) *phon mbin* “white lofty” is to be interpreted as “upper reaches of the White River”. The interpretation of (1) *phon mbin thie tha* as *Bai Shang Da Guo*

“White Upper Great State” is so widespread in China that Bai Shang in the meaning “Tangut State” has even become an entry in standard reference book (Cai Meibiao 1986: 121).

The Japanese scholar Nishida Tatsuo (1964: 161; 1986: 69) obviously supports the “upper reaches of the White River” theory, as is clear from his translation of the name of the Tangut state — *Pai-shang*.

In my opinion, the grammatical mistakes in the translation of the name of the Tangut State are:

1. on the level of the vocabulary — the adjective (5) *mbin* “high, lofty” is translated into Chinese as *shang*, whereas there is no doubt that it has to be translated as *gao*;

2. on the syntactic level — a) adjectives occupying different positions with respect to the noun are translated as sharing one and the same syntactic function, b) adjectives standing in preposition to a noun are regarded as attributes to this noun.

In my opinion, grammatically correct translation of the official name of the Tangut Empire is “The Great Empire of the White and Lofty”.

Now let us turn to the idea concealed in the collocation “white and lofty”, which obviously served as a symbol of the Tangut Empire.

In my work (the first publication of the results of this work is Kepping 1994) I have used the same Tangut material which Nevskij had at his disposal:

1. Tangut 12th century dictionaries kept in the Manuscript Department (Kozlov collection) of the Institute of Oriental Studies in St. Petersburg:

1a. (22) *mi za ngwu ndzie mbu pa ngu nie* “Tangut-Chinese Timely Pearl in the Palm”, usually referred to in scholarly circles by its abbreviated Chinese name *Zhang Zhong Zhu*, published in various editions by Nishida (1966: 186–223), Kwanten (1982), Huang *et al.* (1989);

1b. “Ideographic Miscellany”, referred to by its Chinese name *Za Zi*, prepared for publication by Terentyev-Katansky (forthcoming);

1c. (23) *we ngon* “Sea of Ideograms”, generally referred to as the *Wen Hai*, published by Kepping *et al.* (1969) and by Shi *et al.* (1982). The “Sea of Ideograms” provides a graphic analysis for each Tangut character and its constituents. If the meaning of an ideogram is unclear, additional data can be obtained on the basis of the graphic analysis. Of extreme importance are the groups of homophones which are marked by full stops in the text of the dictionary. A group of homophones runs from one full stop to another. The homophones have one and the same initial consonant and belong to one and the same rhyme.

2. The ode “Hymn to the Sacred Ancestors of the Tangut”, Nevsky (1960: I, 36) translates the word (24) *kia* in the title as “ode”, but the term apparently corresponds to Chinese *ge* and Tibetan *gyer* “ritual song”.

3. The Tangut proverbs (Kychanov 1974).

4. The examples adduced by Nevsky (1960: I, 51) from the Tangut encyclopaedia “Sea of Meanings Established by the Saints”.

5. Tangut poems.

Before setting forth my own view point, I would like to pursue Nevsky's analysis of the tripartite collocation (19) *ma phon mbin* to its logical conclusion to show why the translation "upper reaches of the White River" is unfounded.

First of all, it is necessary to say a few words about the names of the rivers in Tangut. In Tangut texts in the meaning "river" I have never met the word (25) *ma* which is used in the collocation (19) *ma phon mbin*. The meaning "river" in Tangut texts is usually expressed by the word (26) *zie*. In keeping with the Tangut grammar the names of the rivers, nouns, stand in front of the word (26) *zie* "river", e. g. (27) *kin zie* the Jin River, whereas adjectives stand after the word "river", e. g. (28) *zie nia* the Black River (famous *Hei Shui* or Edzina). I have found the word (25) *ma* only as a proper name of a river (noun) in a collocation (29) *ma zie*. This collocation in Tangut translations of Chinese texts stands for the Yellow River (in Chinese *Huang He*). This is corroborated by the use of (25) *ma* in the Tangut expression (30) *tshwu ma* "Milky Way" to be found in the dictionary *Zhang Zhong Zhu*, whereby it is probable that the Tangut shared the view of the ancient Chinese that the Milky Way was the heavenly continuation of the Yellow River in the firmament. It should be noted that in the "Sea of Ideograms", the dictionary entry for the ideogram (25) *ma* first lists the combination (29) *ma zie* as an example, a collocation which I believe meant the Yellow River. Therefore the collocation (19) *ma phon mbin* is to be translated as "The white and lofty Yellow River". I was very much surprised by the fact that for the name of the Yellow River in Tangut, (25) *ma*, the dictionary "Sea of Ideograms" lists sixteen (!) homophones (Kepping *et al.* 1969: 1, 138—9, No. 718—734), one of them (No. 734) being *ma* "mother". Most of the homophones are connected with the meaning "female", "mother", but there are also real homophones, e. g. No. 730 (62) *ma* "presumably". It seems that the ideogram (25) *ma* in the Tangut name of the Yellow River symbolizes the river as the embodiment of the female principle, i. e. "Mother River". Consequently, the collocation (19) *ma phon mbin* "the white and lofty Yellow River" and (31) *ma phon mbin* "the white and lofty Mother" are homophonous. However, only a part of the collocation (19) *ma phon mbin* serves as a symbol of the Tangut Empire and is included into the official name of the Tangut Empire, namely, (12) *phon mbin* "white, lofty". I believe that this is a fixed collocation which expresses the idea of the union of the female and male principles, since (5) *mbin* has a homophone (32) *mbin* "membrum virile" (Kepping *et al.* 1969: 2, 87, No. 3981), whereas (4) *phon* can easily stand for the female principle (in Tangut the sun, which represents the female principle, has a fixed attribute "white"; for details see below).

Corroboration for the hypothesis that the collocation (12) *phon mbin* symbolizes the union of the female and male principles, corresponding to Tibetan *yab* and *yum*, is found in the sense in which (4) *phon* and (5) *mbin* are used on a Tangut icon on a display in the Hermitage (Rudova 1993: 206). The icon which represents Guan-yin (= Avolokiteshvara) bears Chinese captions alongside two figures of donors in the bottom right-hand corner (fig. 3). These two female figures in rich gold-embroidered red robes evidently belong to the emperor's family — Tangut laws prohibited gold-embroidered robes even for the emperor's

relatives (Kychanov 1987b: 223). They differ in age, the woman on the left is obviously older than the one on the right. The similarity of the two women suggests that they are mother and daughter. The cartouche to the left of the elder woman bears the caption *Bai shi tao hua* "Peach flowers from Madame Bai", and to the left of the younger woman the caption is *sin fu Gao shi yin jiang xiang* "Incense from the newly wed Madame Gao" [3]. I presume that this is an allegory: the elder lady — Madame Bai — is the Tangut empress, and as such, personifies the female principle *bai*, "white", whereas the younger lady — Madame Gao — is her daughter, the princess, bearing her father's (the emperor's) symbolic surname *Gao*, representing the male principle. In other words, in Tangut culture the female and male principles were personified by the emperor and empress respectively. It seems that the characters *bai* and *gao* were taboo and could not be used as the surnames for the common people, since among the Chinese family names listed in the above-mentioned Tangut dictionary "Ideographic Miscellany" in the section (33) *za mie me* "Chinese surnames", the names (4) *phon* (= *Bai*) and (5) *mbin* (= *Gao*) are not listed, and corresponding Chinese names are rendered by other characters — (34) *phe* for *Bai* and (35) *keu* for *Gao*.

The Tangut collection held at the British Library contains a fragment of an imperial copy of a manuscript, call number OR 12380 (3302), written on yellow paper, as was customary for both the Chinese and Tangut cultures for personal copies of manuscripts belonging to the emperor. Both sides of the manuscript bear a fragment of the 21st juan (scroll) of the Mahāparinirvāṇasūtra. In the middle of the manuscript, on a background especially darkened with imperial ochre appear the emphatically enlarged ideograms (5) *mbin* and (36) *tshia* "virtue" immediately followed by the name (37) *nin po tsa* "Emperor Bodhisattva", the last three characters being smaller than the first two. The "improper" position of the adjective (5) *mbin* "lofty" gives here a clear hint that what was actually intended is its homophone (32) *mbin* "membrum virile" in the function of an attribute to the following noun. I believe that the five-character combination (38) *mbin tsia nin po tsa* means "Emperor Bodhisattva [possessing] the VIRTUE OF MBIN". It seems that the Tangut emperors were regarded as "sacred kings". If so, then we are now faced with a challenge of trying to reconstruct the rituals concerning Tangut sacred kings recorded in native Tangut texts.

The study of Tangut poetry and proverbs has demonstrated that the Tangut believed that the surrounding world at all levels (cosmos, nature and man) reflected the coalescence of the female and male principles. In the proverbs the sun is always connected with the female principle, the moon with the male principle. In the following proverb the mother is directly compared to the sun, which bears the female epithet "white", and the father is compared to the moon, which bears the male (?) epithet "red":

(39) <i>tu</i>	<i>mbe</i>	<i>phon</i>	<i>su</i>	<i>ma</i>	<i>ldiu</i>	<i>ldwon</i>
thousand	sun	white	compare	mother	beauty	hot
<i>khi</i>	<i>lhie</i>	<i>nin</i>	<i>su</i>	<i>via</i>	<i>sie</i>	<i>sweu</i>
ten thousand	moon	red	compare	father	wisdom	bright

(Proverb No. 8)

A mother's beauty is hotter than a thousand white sun,
A father's wisdom is brighter than ten thousand red moons.



Fig. 3

From the proverbs it becomes clear that the epithet "white" for the sun and the "red" for the moon were apparently fixed collocations: see for example proverb No. 199 (Kychanov 1974). Similarly (40) *me* "heaven" is often compared with (41) *ldie* "earth", whereby heaven apparently represented the male principle and the earth the female principle. The term (41) *ldie* "earth" has a synonym (42) *ma* "earth", homophonous to (43) *ma* "mother". In the following proverb (40) *me* "heaven" bears the "male" epithet (5) *mbin* "lofty, high":

(44) <i>me</i>	<i>mbin</i>	<i>mi</i>	<i>ton</i>	<i>ni</i>	<i>mbi</i>
heaven	high	not	help	clouds	low
<i>ldie</i>	?	<i>mi</i>	<i>ton</i>	<i>tsia</i>	<i>ru</i>
land	broad	not	help	road	narrow

(Proverb No. 335)

A high heaven will not help if the clouds are low
Extensive lands are of no use if the road is narrow [4].

In the Tangut ritual songs the opposition of (45) *lie lan nge* "in the west are the Helan mountains" vs. (46) *vie ma nio* "in the east is the Yellow River" recurs again and again. The Yellow River, which, as we have established, is a female entity, is associated with the east, while the Helan mountains, embodying presumably the male principle, are linked to the west. It should be noted that the word (47) *nio* "river" in the collocation (46) *vie ma nio* "in the east is the Yellow River" is a homophone of the words (48) *nio* "east" and (49) *nio* "source". The following table shows pairs of nouns which are regularly opposed to each other in Tangut texts as female and male personifications:

female (4) <i>phon</i> "white"	male (5) <i>mbin</i> "lofty"
sun	moon
earth	heaven
east	west
Yellow River	Helan mountains
any river	any mountain
empress	emperor

I suppose that a central, prevailing Tangut notion is that the Tangut State was born from the merging of the lofty Helan mountains in the west, the male principle, and the white waters of the Yellow River in the east, the female principle. When for the first time I have visited the territory which previously belonged to the Tangut Empire I was deeply impressed by the two natural dominants in the landscape — the mighty flow of the Yellow River and the impressive beauty of the Helan mountains. And — it is of importance for this study — the tombs of Tangut emperors are located at the foot of the Helan mountains. This is really an unforgettable view: one can easily imagine the influence and significance which the Yellow River and the Helan mountains had in the everyday life of the Tangut eight centuries ago.

Now let us turn again to the collocation (19) *ma phon mbin* "the white and lofty Yellow River" which is homophonous to "the white and lofty Mother".

The usage of this collocation in native Tangut texts shows that most likely this collocation denotes a mountain, not a river. This meaning is evident from the adduced below example from the encyclopaedia "The Sea of Mean-

ings Established by the Saints". The text of the encyclopaedia is arranged in vertical lines, wherein the basic text consists of segments of four to six ideograms, each segment immediately followed by a trailing commentary written in slightly smaller ideograms. This example is taken from the section entitled "Names of Mountains":

(50) <i>nge</i>	<i>kwe</i>	<i>wa</i>	<i>ndzio</i>	<i>viei</i>	<i>nge</i>	<i>zie</i>	<i>ndzio</i>
mountain	body	broad	tall	snow	mountain	all	tall
<i>mi</i>	<i>mba</i>	<i>riu</i>	<i>lhie</i>	<i>zi</i>	<i>ni</i>	<i>me</i>	<i>ion</i>
not	interrupt	all	state	all	reach	root	source
<i>ma</i>		<i>phon</i>	<i>mbin</i>	<i>ngu</i>			
Yellow River(=	mother)	white	high	is			

The mountain ridge is wide and long. [Commentary:] The snow-capped mountains are high and long. They go on continuously and reach all countries. The place where [our country] begins is the "White and Lofty Mother [mountain]".

Since the entire section "Names of Mountains" is devoted specifically to alpine toponyms, it seems obvious to assume that (19) *ma phon mbin* denoted a mountain within the Tangut realm, most likely in the snow-capped Helan mountain ridge. The assertion that (19) *ma phon mbin* is not a river toponym is corroborated by the fact that this name is lacking in the section (51) *ma ngon* "Rivers and Seas" of the dictionary "Ideographic Miscellany" (Terentyev-Katansky, forthcoming), whereas (29) *ma zie*, the name for the Yellow River, is included. Furthermore, graphic analysis of the ideogram (52) *so* "high, tall", which is an adjective often found in contexts where snow-capped mountains are specifically involved, provides further support for our assertion. The ideogram (52) *so* consists of the left portion of the ideogram (4) *phon* "white", representing the female principle, and of the entire ideogram (5) *mbin* "lofty, high", representing the male principle. Graphically, therefore, the ideogram (52) *so* symbolizes the union of the female and male principles. The adjective (52) *so* "high, tall" is used in the following proverb:

(53) <i>nge</i>	<i>kha</i>	<i>vie</i>	<i>tiei</i>	<i>ta</i>	<i>so</i>
mountain	on	snow	lie	topic	high
<i>mie</i>	<i>ngu</i>	<i>tshia</i>	<i>ndziei</i>	<i>ta</i>	<i>piu</i>
person	among	virtue	possess	topic	respect

(Proverb No. 297)

If there is snow on the mountain, the mountain is tall
If a man has a sense of dignity, he is respected.

The next example, this time from the Tangut poems, is another prove for the assertion that (19) *ma phon mbin* means a mountain:

(54) <i>ma</i>	<i>phon</i>	<i>mbin</i>	<i>me</i>	<i>u</i>	<i>ldie</i>	<i>mie</i>
Yellow River	white	lofty	heaven	head	earth	tail
(= mother)						

The head of the White and Lofty Mother is [in] the heaven,
[her] tail is [in] the earth.

The earliest (A.D. 1346) Tibetan source on the Tangut (*Hung Shi*) begins the description of the Tangut state mentioning a mountain somewhere near Liangzhou (modern Wuwei in Gansu province). This mountain is a residence of

a spirit [5]. The sacred Tangut mountain reflected in the Tibetan texts was known to R. A. Stein (1947—1950) as early as in the 1940's. Can this mountain be identified with the mountain "The White and Lofty Mother"? Today I have no answer to this question.

Now let us return to the second line of the "Hymn to the Sacred Ancestors of the Tangut" (see (17)) which Nevsky (1960: 1, 38) translated as "the paternal mounds of the red-faced ones in the upper reaches of the White River". The word (55) *lhiuo* used in this line is usually translated as "grave". Yet in the texts I have studied another word stands for "grave", viz. (56) *lwa*. Graphically the ideogram (55) *lhiuo*, according to the native Tangut dictionary "Sea of Ideograms", consists of parts of the ideograms (52) *so* "tall", (57) *ndziu* "ten" and (58) *rie* "skillful", and the contexts in which it is used indicate that (55) *lhiuo* should be translated specifically as "emperor's tomb" or "imperial grave". Hence, the following two conclusions can be drawn: 1) "the paternal burial mounds of the red-faced ones" are the graves of Tangut emperors; 2) the Tangut emperors were believed to have been the fathers of "the red-faced ones". The first conclusion is important to the present study. The ramifications of the second conclusion is dealt with in Kepping (forthcoming). The second line of the "Hymn to the Sacred Ancestors of the Tangut" should therefore be translated as "the tombs of emperors, fathers of the red-faced ones, are located near [the mountain called] the "White and Lofty Mother"". The tombs of Tangut emperors, as we know, are located at the foot of Helan mountains. And apparently the mountain the "White and Lofty Mother" is somewhere in Helan mountains as well. Hence the "White and Lofty Mother" was the name of a peak in Helan mountains and everybody understood

the name of the Tangut State in terms of metonymy, named after a geographical landmark, much in keeping with the traditions of the neighbours (cf. *Liao guo* was named after the *Liao* River). However, the actual significance of the name of the Tangut Empire becomes evident only at the level of written language, i. e. script itself. In the name of the mountain instead of the ideogram (43) *ma* "mother", its homonym (25) *ma* "Yellow River" was used. To demonstrate the inseparable union of the female and male principles a mountain (personification of a male principle) was called a river (personification of a female principle).

In conclusion I would like to stress the significance of native Tangut texts not only for the reconstruction of the complex of knowledge about Tangut Empire which is called in Tangut ritual songs (61) *khi ndvu lhie* "the State of ten thousand secrets", but at the same time because these texts reveal some deep-related notions which played a significant part in the relations between the so-called dynasties of conquest (Tangut, Khitan, Jurchen, Mongol). Perhaps these notions were shared by all these dynasties as well. If I am correct in my interpretation of the official name of the Tangut State as a mountain, that means that Tangut Empire itself personified the male principle. Then we have to assume that the Liao State, called after the Liao River, personified the female principle. As to the Jin (=Jurchen) dynasty, the Cambridge history of China (1994: 221) states that it was named after the An-chu-hu River, which means that the Jin State has personified the female principle as well. No doubt these mythological notions were decisive not only at the level of the marriage alliances, but shaped the whole civilization of the dynasties of conquest [6].

List of Characters

1. 𐰇𐰏𐰣𐰚𐰣 2. 𐰇𐰏𐰣𐰚𐰣 3. 𐰇𐰏 4.
- 𐰇𐰏 5. 𐰇𐰏 6. 𐰇𐰏 7. 𐰇𐰏 8. 𐰇𐰏𐰣𐰚 9. 𐰇𐰏𐰣𐰚 𐰇𐰏
10. 𐰇𐰏𐰣𐰚 11. 𐰇𐰏𐰣𐰚 12. 𐰇𐰏𐰣𐰚 13. 𐰇𐰏 14. 𐰇𐰏 15.
- 𐰇𐰏 16. 𐰇𐰏𐰣𐰚 17. 𐰇𐰏𐰣𐰚 𐰇𐰏 18. 𐰇𐰏𐰣𐰚 𐰇𐰏 19.
- 𐰇𐰏𐰣𐰚 20. 𐰇𐰏𐰣𐰚 21. 𐰇𐰏𐰣𐰚 22. 𐰇𐰏 23. 𐰇𐰏 24. 𐰇𐰏 25. 𐰇𐰏
26. 𐰇𐰏 27. 𐰇𐰏 28. 𐰇𐰏 29. 𐰇𐰏 30. 𐰇𐰏 31. 𐰇𐰏 32. 𐰇𐰏
33. 𐰇𐰏 34. 𐰇𐰏 35. 𐰇𐰏
36. 𐰇𐰏 37. 𐰇𐰏 38. 𐰇𐰏 39. 𐰇𐰏
40. 𐰇𐰏 41. 𐰇𐰏 42. 𐰇𐰏 43. 𐰇𐰏 44. 𐰇𐰏 45. 𐰇𐰏

- Nevskii N. A., 1960 [posthumous]. *Tangutskaia filologija* (Tangut Philology) (Moscow).
- Nishida Tatsuo, 1964/1966. *Seikago no Kenkyū, Seikago no Saikōsei to Seikamoji no Kaido-ku* (A Study of the Hsi-Hsia Language, Reconstruction of the Hsi-Hsia Language and Decipherment of the Hsi-Hsia script), i—ii (Tokyo).
- Rudova M. L. "Guanyin", *The State Hermitage Museum. Lost Empire of the Silk Road*.
- Shi Jinbo, 1988. *Xixia Fojiao Shi Lue* (Yinchuan: Ningxia Renmin Chubanshe).
- Stein R. A., 1947—1950. "Mi-nag et Si-hia. Géographie historique et légendes ancestrales", *Bulletin de l'Ecole Française d'Extrême-Orient*, 44, pp. 223—65.
- Terent'ev-Katanskiĭ A. P. (forthcoming). *Tangutskii slovar' "Ideograficheskaja Smes'"* ("The Tangut Dictionary "Ideographical Miscellany"").
- Wang Jinru, 1932. *Xixia Yanjiu, Di Yi Ji* (Beiping: Guoli Zhongyang Yanjiu Yuan Lishi Yuyan Yanjiu Suo).

Illustrations

- Fig. 1. The preface to the Tangut translation of the Buddhist canon (Stockholm Ethnographical Museum, Courtesy of the trustees of the Sven Hedin collection).
- Fig. 2. Fragment of the Chinese MS containing the sūtra about bodhisatva Sheng-i (St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, TK-145).
- Fig. 3. 12th century scroll on silk (Courtesy of the State Hermitage Museum, St. Petersburg).

PRESENTING THE COLLECTIONS

T. A. Pang

RARE MANCHU MANUSCRIPTS FROM THE COLLECTION OF THE ST. PETERSBURG BRANCH OF THE INSTITUTE OF ORIENTAL STUDIES, RUSSIAN ACADEMY OF SCIENCES*

To the history of the collection

The Manchu collection of the Institute of Oriental studies, St. Petersburg, is the largest outside China. The printed catalogues of this collection number up to 600 entries of manuscripts and block prints [1]. The first Manchu documents were brought to St. Petersburg in the beginning of the 18th century, but the actual study of Manchu was started much earlier with the establishment of commercial, and later, diplomatic relations between Russia and the Qing dynasty China in the middle of the 17th century. At that time all official Qing documents were written in Chinese, Mongolian and Manchu, and Manchu was used during diplomatic negotiations.

The members of the Russian Ecclesiastic mission in Peking were the first to collect and bring to Russia the Manchu manuscripts and block prints. They were the pioneers of Manchu studies to compile the first Manchu-Russian dictionaries, grammars and readers, as well as to make a linguistic and literary research [2]. Many of these materials remained unpublished and are kept in the Archives of Orientalists.

Different Manchu grammars, dictionaries, historical and philosophical texts both in manuscripts and block prints were brought to St. Petersburg Asiatic Department of the Russian Ministry of Foreign Affairs, and later to the Asiatic Museum of the Imperial Academy of Sciences. Many of them are well represented in the world libraries owing to the collections of sinologists and Jesuits from different countries. Nevertheless, the Manuscript Department and the Archives of Orientalists contain a unique Manchu material. This fact allows to put the Institute collection on the first place among the other depositories both in number and repertoire of the Manchu manuscripts. Many of them were acquired from the outstanding man-churists N. N. Krotkov and A. V. Grebenshchikov, who worked on the Institute collection at the beginning of this century. The Institute archives could be very handy to reconstruct their scientific activity.

Nikolai Nikolaevich Krotkov (1869—1919) graduated the Manchu-Chinese department of the Oriental faculty, St. Petersburg University. In 1894 he started his diplomatic career at the Russian consulate in Kuldzha as the head of the consulate school of translators and interpreters. Later he was a secretary at Russian consulates in Jilin, Qiqihar and Kuldzha, and then a Russian consul in Urumci, which he left for Russia in 1912 [3]. In spite of his diplomatic activity, N. N. Krotkov continued his studies of the Orient. Thanks to him the St. Petersburg Asiatic Museum got one of the largest collection of Central-Asiatic documents like Uighur, Indian, Mongolian and Manchu manuscripts and block prints. He sent the first manuscripts in 1898 [4], his Turfan and Central-Asian materials are mentioned in the letters of S. Th. Oldenburg and V. V. Radlov [5]. In 1910 N. N. Krotkov became a correspondent of the Anthropological and Ethnographic Museum and was given an honorary title of a correspondent member of the Russian committee for the studies of Middle and East Asia [6]. Of his personal interest were Manchu studies. Therefore he had collected a lot of precious Manchu texts which were later added to the Institute collection [7]. On his return to Petrograd in 1918 N. N. Krotkov compiled a catalogue of Manchu books kept in the Asiatic Museum listing 323 entries of manuscripts and block prints [8]. This list of Manchu books did not include the manuscripts from his private collection. They were listed later by M. P. Volkova and K. S. Iakhontov [9]. The Manchu texts of N. N. Krotkov are a unique registration of the spoken Sibe dialect of Manchu as he heard it at the merge of the centuries. They mostly include epic songs and shamanic texts. To collect them he asked for the help of his Sibe friends in Kuldzha, Suiding, Huiyuancheng and Chuguchak. Some of their names could be reconstructed from the letters in his personal file in the Archives of Orientalists.

The other private collection of Manchu materials belonged to the famous Russian manchurist A. V. Gre-

* The present article is an enlarged English version of the paper read in Chinese at the conference on Manchu studies in Peking, 1992 and published in *Manxue yanju*, No. 2, Peking 1994, pp. 195—200.

benshchikov (1880—1941). The materials on his scholarly and research activity are well presented in the Archives of Orientalists either (file 72). It is known that A. V. Grebenshchikov entered the Oriental Institute, Vladivostok, as a voluntary student in 1902, as a regular student of the 4th year in 1906. Next year, in 1907, he graduated the Chinese-Manchu department and remained for teaching there [10]. In 1907—1909 A. V. Grebenshchikov studied the Tungus-Manchu linguistics and phonetics in particular. During this period he made his first trips to Manchuria collecting the material for his master thesis on Manchu shamanism. From 1918 he worked as a professor at different chairs of the Vladivostok Oriental Institute and the Far-Eastern Branch of the USSR Academy of Sciences. He published over 50 works on the Manchu and Chinese languages, culture and literature [11]. Almost all of them were written on the materials collected by the scholar during his ethno-linguistic expeditions to Manchuria in 1908—1927 [12]. In 1935 A. V. Grebenshchikov moved to Leningrad and started his work in the Institute of Oriental Studies, USSR Academy of Sciences. He put his task to describe the Institute Manchu collection, make research on the Manchu shamanism and translate into Russian "*Wecece metere kooli bihe*" ("A Book on Shamanic Rights"). He finished his translation in 1939, but could not publish it [13] because of the World War II. A. V. Grebenshchikov died in the besieged Leningrad on October 15, 1941 [14].

N. N. Krotkov and A. V. Grebenshchikov worked in the Manuscript department of the Institute studying the

Manchu collection. It was only in 1965 when the catalogue of Manchu manuscripts was published by M. P. Volkova who based her research on the work of these outstanding scholars, as well as on the notes done by V. L. Kotvich.

For many years M. P. Volkova was the head of the Manuscript department leading her scholarly research in Manchu studies. She translated into Russian and published the Manchu manuscript from A. V. Grebenshchikov's collection — "*Nisan samani bihe*" [15]. Publication of this unique text pushed a discussion on the existence of the Manchu original literature and peculiarities of Manchu interpretations of the Chinese classical and historical texts. The text from the St. Petersburg collection has already been translated into nine languages and other versions of this text were published [16].

The compilation of Manchu catalogues started by M. P. Volkova with the "Description of Manchu Manuscripts from the Institute of the Peoples of Asia, USSR Academy of Sciences" in 1965 resumed the studies of Manchu collections all over the world. Since that time more than 30 catalogues of the main libraries of Europe, America, China, Japan, India and Mongolia have been published. The material registered in them gives a more exact idea of the importance of the Institute collection. The Manuscript department of the Institute of Oriental Studies, St. Petersburg, along with the standard repertoire contains unique original Manchu texts and copies. Some of the most important I am going to introduce below.

History of culture

The Institute Manchu collection is rich with historical and philosophical texts, most of them being translations from Chinese. Nevertheless, it includes a rare example of the Manchu original literature "*Emu tanggu orin sakda i gisun sarkian*" ("Stories, Told by One Hundred and Twenty Sages") (call number B 15, see *fig. 1*). It consists of one hundred and twenty stories on Manchu history and culture. It was written by Sun Yun (1752—1835), a Mongol from Marat clan [17], in 1789 when he was on service in Urga. In 1791 the text was revised by Furengtai, a Manchu from Wangyan clan. In 1809 Fu-jun (1749—1834), a Mongol from Jot clan, translated it into Chinese. This history of the text is registered in the prefaces of these three men. The Institute manuscript belonged to Fu-jun and bears his personal red seals at the end of his preface

(f. 10a) as well as at the end of each of the eight notebooks. It seems, this is the only existing Manchu-Chinese copy of the text. All the other known texts kept in the libraries of America, Japan, Mongolia and China are written in Manchu and do not have Fu-jun's preface. The most complete Manchu text from Chicago library was published in 1982 by the Centre for Chinese documents in Taipei, and was translated into German in 1983 [18]. The Institute manuscript differs from the other known copies by the additional preface of Fu-jun, Chinese translation, the ways of arrangement of the chapters as well as lexical peculiarities. Therefore, linguistic and textological study of the existing variants in comparison with the Institute's bi-lingual text could be of great interest.

Shamanism

Almost all the manuscripts on Manchu shamanism were collected by A. V. Grebenshchikov and N. N. Krotkov and date back to the end of the 19th—beginning of the 20th century. It was the period of great interest to ethnography and folklore of minor peoples. At that time Russians could easily work in Manchuria contributing a lot of ethnographic material to the Russian collections. The shamanic literature usually existed in oral tradition; the texts were transmitted from the shaman-teacher to his apprentice. The Russian scholars were the first to discover the scientific value of shamanic prayers. They asked their informants to write down the Manchu texts, or themselves

transcribed these texts in Russian transliteration [19]. We have already mentioned the famous text of A. V. Grebenshchikov "*Nisan saman-i bihe*". Its publication by M. P. Volkova in 1961 initiated scientific studies of this manuscript, its versions and variants. At present the Daghur, Evenk and other variants of this text are found. This proves the popularity of this subject among the other Tungus-Manchu peoples [20].

All of the Grebenshchikov manuscripts were written down in Manchu villages during his trips to the north-eastern China in 1908—1927. Some of them are followed by drawings of the shaman costumes, drums, musical no-

tes. Of the most interest is his notebook (file 75, opis' 1, No. 47), which includes "*Wecece juktere sain ehebe tuwaranje bihe*" ("A Book on Good and Evil during the Sacrifices"), "*Targacun tolon ehe sain juwan ilmun han*" ("A List of the Ten Evil and Good Masters of the Hell") [21], "*Siyoo wen di juktehe-i kumun maksu be tok-tobuha wasimbuha selgeri hese*" ("Highly Approved Regulations on Singing and Dancing during Sacrifices in the *Xiao-wen-di* Temple" (fig. 2). The notebook No. 48 contains everyday stories collected by A. V. Grebenshchikov in Aigun and Hulguri during his stay there from August 22 till September 1, 1917. Alongside with these stories we find there "*Weceku gisun*" ("Words Pronounced during Sacrifices").

The Manchu collection of the Manuscript department has a curious album "*Sula i bihe*" ("Trifles") (call number A 156), which contains coloured drawings by A. V. Grebenshchikov of the Manchu's life (fig. 3). The expedition notes of Grebenshchikov bear an ethnographic interest depicting the Manchu everyday life in the north-eastern China at the beginning of the 20th century.

Another important text for the studies of Manchu shamanism, which follows "*Nisan saman-i bihe*", is the manuscript from N. N. Krotkov's collection "*Saman küwaran i bihe*" ("Shaman Court") (call number B 55). Prof. Gimm referred to it as one of the most interesting and difficult for understanding [22]. The first manuscript tells about the shamaness' journey to the underworld in search of the dead young man's soul and gives a description of the shamanic rite of resurrection of the soul. The second text is a collection of different shaman prayers and invocations.

As it was mentioned above, N. N. Krotkov had collected many precious materials in Xinjiang, the most important being "*Saman küwaran i bihe*" copied on his request. Among the documents in N. N. Krotkov's file there is a manuscript on the Manchu wedding ceremony (file 32, opis' 1, No. 8, folio 16) written in the same handwriting with a sharp horizontal ending of the last letters. It corre-

sponds to the letter of a Manchu bannerman Enden to Krotkov written on a small piece of paper (file 32, opis' 2, No. 5). We may assume that both of these texts were written down by Enden. He could not be the author or a compiler of "*Saman küwaran i bihe*", since the text itself twice gives the date as "the first day of the tenth month of the third year of *Badarangga doru*", i. e. November 5, 1877. It is 20 years after N. N. Krotkov's arrival to China. In the "Description" of M. P. Volkova the date of the text is by mistake given as 1882 (No. 59). Nine prayers of several shamans are registered in the manuscript. The first one was read by a shaman apprentice, who was born in the year of monkey and belonged to the Gejuru clan. His teachers were from the Gijoro and Gurgiya clans. The second prayer belonged to Gesguri, a shaman apprentice from the Hashuri clan, who was born in the same year. The third shaman was Karanju from the famous Gejer clan.

The manuscript is a compilation of different prayers and invocations. The first two were recited by the shaman apprentice during his initiation and includes invocations to the spirits for helping him in ceremonies. The following three texts are the prayers to the spirits *Mafa* and *Mama* (male and female ancestors). Three other prayers were read during healing practice. They illustrate shaman's journey through different underworld villages guarded by different spirits, different obstacles, and preparation of ritual objects used during expel of the illness from a sick body. The last prayer of the shaman Karanju tells about his journey to the highest deity *Isanju-mama*, who is imagined to be sitting on the top of the golden ladder and who patronizes all shamans and their help-spirits [23].

"*Saman küwaran i bihe*" is important for being the variant of the other Sibe-Manchu manuscript "*Saman jarin*" which was published in Urumci in 1990. It has been already translated into Chinese and German [24]. Both manuscripts are an important and valuable source for the studies of Sibe-Manchu shamanism and the Sibe-Manchu language of the last century [25].

Folklore

The third largest part of the Manchu collection consists of the Sibe-Manchu poetry and folklore. The most of them is dedicated to the Sibe transfer from Manchuria to Xinjiang in the second half of the 18th century. Recently several poems on this event were published in China [26]. To these texts we may add two variants from our collection: "*Ba na i ucun*" (call number A 51) — "Song about homeland" (fig. 4) and "*Mukden ci tucike de ba i ucun*" (call number A 29) ("Song on the Transfer from Mukden"). If compared, the second poem turns to be an abridged variant of the first text. It lacks the first part which glorifies the Qing dynasty, the conquer of the Jungar and establishment of the western borders of the Qing period China. It also speaks on the cleverness of the imperial laws and the imperial order on the transfer of the Mukden Sibe to the basin of the Ili river. Both texts coincide from the folio 2b of "*Ba na i ucun*", still having textological and poetical differences. Both variants are very important for the history of Manchu literature, and along with already published variants allow to trace the development of this text from a simple folk song to a big poetical composition during hundred years of its existence. Both of our manu-

scripts were translated into German and published in 1988 by Prof. G. Stary in "Epic Songs of the Sibe-Manchu" [27].

The same could be addressed to another text "The Song on Kashgar", dedicated to the Manchu fighting against the Kashgar uprising in 1820—1828. The Chinese and Sibe scholars have published several versions of this texts [28]. As it is in the previous case, the earliest versions are kept in the manuscript collection of the St. Petersburg Branch of the Institute of Oriental Studies. The first short manuscript was registered by V. V. Radlov and published later by A. O. Ivanovsky in 1895 [29]. The second manuscript was copied on the request of N. N. Krotkov in 1903 and is kept under a call number A 16, titled "*Daicing guruni kasigar ucuni bihe*" (fig. 5). This is the most complete version of this text, and the comparative study of the existing versions will give a history of the development of it.

The third worth mentioning text from the Institute's collection is "*Muduri mukdehe uculen*" (call number A 53) ("Song on the Rise of the Dragon") [30]. It glorifies the Qing dynasty, sage Manchu emperors and courage of the bannermen. The Manchu text on the top of the page is

咸登壽域用協
 祇期薄海內外人民

諭旨將時憲書花中之數
 疊算重週
 增為百二

皇上持降

恭維

百二老人語序

Handwritten vertical text in cursive script, likely the main body of the document.

Fig. 1

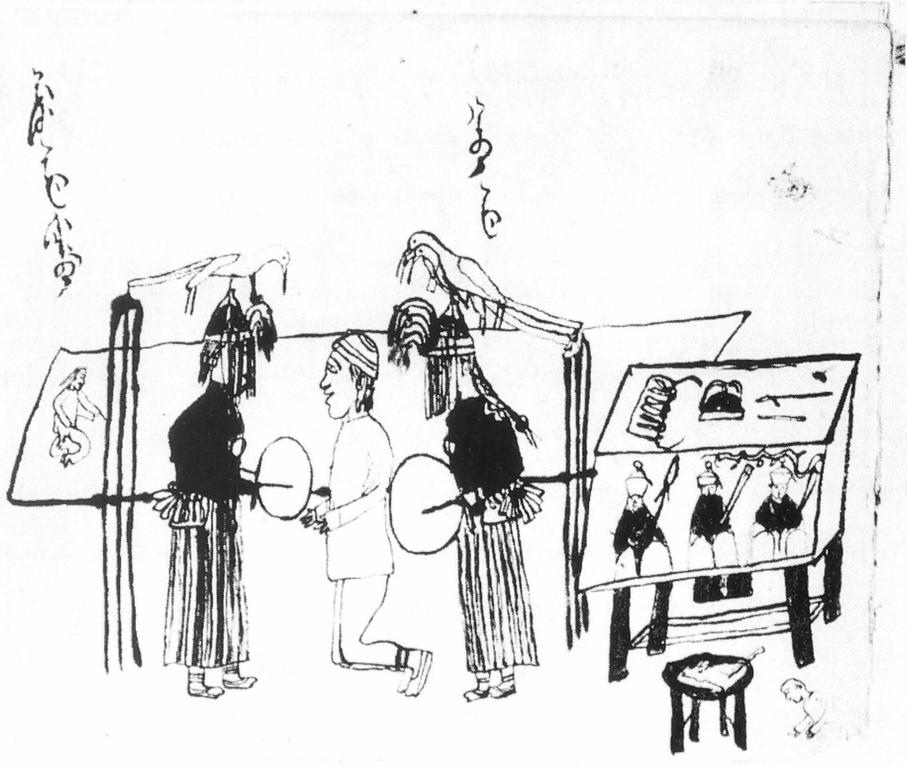


Fig. 3



The image shows a page of a Manchu manuscript with five vertical columns of text. The script is Manchu, written in a cursive style. The text is arranged in five columns, reading from right to left. Each column contains several lines of text, with some characters appearing to be repeated or forming a specific sequence. The paper is aged and slightly yellowed.

Fig. 4



Fig. 6

translated into Chinese, but this Chinese text at the bottom of the page is written in Manchu alphabet (fig. 6). This kind of transliteration of the Chinese texts is often registered in the early Manchu documents. It enables the Manchus to be aware of the Chinese pronunciation in case they did not know the hieroglyphs.

The Manchu collection includes special occasion or festivity songs like *sarin ucun*. They were sung during wedding parties by the newly-weds, match-makers and guests. Some of them were already published in Chinese [31], but the Institute's "*Sarin de uculere ucun*" (call number A 6) ("Song Which is Sung during the Party") is the most complete and poetical. It consists of 57 stanzas, 4 lines each. The other popular song from the St. Petersburg collection is "*Fadu seolere ucuni mudan*"

(call number A 3) which was sung by a young lady when she was embroidering a purse for her beloved before his departure. The text was written down for N. N. Krotkov on September 26, 1899 (*Badarangga doru orin sunja aniya jakun biya orin juwe de*). In the "Description" of M. P. Volkova this date is by mistake translated as August 8, 1904 (No. 179). The Manchu text is followed by the Chinese translation in Manchu alphabetical transliteration.

All these manuscripts were collected at the end of the 19th—beginning of the 20th century and are the earliest registrations of the Sibe-Manchu folklore texts. The modern versions of some of them were recently published in Chinese [32].

Ceremonies

There are some interesting texts referring to the traditions of the Manchus. The most complete is "*Manju gūšai boo hanjire niyaman jafara dorolon i bithe*" (call number A 116) ("Book of Wedding Ceremonies of the Manchu Banner Families"), which was also brought by N. N. Krotkov. It describes the Manchu engagement, wedding ceremonies of the Sibe-Manchu in Xinjiang. It contains the rules of filial piety, remembrance of the passed parents, as well as the list of terms of family relations. This text was translated into English with transliteration and facsimile [33].

Another similar text is found in the personal file of N. N. Krotkov with other version of this ceremony which was recently published in facsimile with German translation [34].

The descriptions of the Manchu ceremonies give an interesting material on the Chinese influence on everyday life, as well as on preservation of the original Manchu tradition. These manuscripts may interest ethnologists and researchers in the Manchu traditional culture.

Translations of the Chinese literature

Alongside with the original Manchu texts discussed above, the Manuscript collection includes the Manchu translations of the Chinese literature. Most of them are historical and classical texts. The first Manchu translations were done in Mukden in 1632—1633 on war strategy, law and history to enable the reader to get acquainted with the Chinese traditional science and literature [35]. Together with the well known translations of "*San guo zhi*", "*Xi xiang ji*", "*Jing Ping Mei*", that are available in almost every Manchu collection both in block-print and manuscript form, the Institute collection has some Manchu manuscripts of the Chinese novels which do not exist in block-print and their original Chinese version. Many of them were burnt down during the literary inquisitions of Qianlong time (1736—1796) on the pretext of bearing danger to the state and moral order of the Qing empire. Thus, the existing unique Manchu manuscripts give us the idea of the destroyed Chinese texts. Prof. Gimm has numbered 76 titles of such Manchu translations, 22 of these novels and stories are kept in the Institute's collection [36]. 11 of them are not presented in other libraries of the world. Thus, according to the number of the Chinese literary texts in Manchu translation our collection could be considered to be the largest. The earliest manuscript is the novel "*Zui pu ti-i bithe*" (call number B 53) dating back to 1684—

1685. The other manuscript "*Hūwa tu yuwan-i bithe*" was written in 1690. Both novels were translated by a person whose pen-name was *Tian-hua-zang zhu ren* ("The owner of the heavenly flowers chamber"), who presumably was Xu Zhen being known for his nine translations of the Chinese novels done by him in the second half of the 17th century [37]. Since these Manchu translations have not yet been studied we limit ourselves to the list of the other nine Chinese novels in their Manchu translation, which are kept in the Manuscript department:

1. *Can jen giyoo ši-i bithe* (call number B 9);
2. *Jang ze mei-ni bithe* (call number C 5);
3. *Gin yün kiyoo-i bithe* (call number B 31);
4. *Feng huwang c'i bithe* (call number A 97);
5. *Šeng hūwa meng ni bithe* (call number B 7);
6. *San guo yin* (call number A 9), the Russian title on the cover is *Predanie o dobrodetel'nom Syma Mou* ("A Legend about a Good Sima Mou");
7. *Jeo gurun-i bithe* (call number A 62);
8. *Dergi Han gurun-i bithe* (call number A 77);
9. *Ing Liyei juwan-i bithe* (call number B 53).

The study of different translations of Chinese novels and short stories which do not exist in their original version enriches our knowledge of the Chinese literature and literary contacts between the Far-Eastern countries [38].

Newspapers

In addition to these rare manuscripts the Institute possesses a collection of a Manchu newspaper "*Ice donjin*" (call number C 1) published in Hailar in 1925. It contains information about home and international affairs of that

time, as well as historical sketches. While checking this file, we found 20 issues of this paper, the "Description" of M. P. Volkova registers only 17. It turns out that the Institute collection is the biggest, since we know nothing about

Manchu newspapers in the Chinese libraries. In the Library of the Hungarian Academy of Sciences there are several issues of the newspaper "Hulun buir ice donjin afaha" of 1930 [39].

In 1913 A. V. Grebenshchikov published one issue of "Sahaliyan ula erindari boolara hoošan" of 1912 in his "Man'chzhurskaia khrestomatia" [40].

These newspapers could be valuable for the study of the history of Manchu speaking people of that time.

The Manchu manuscripts presented above, from our point of view, are the most interesting in the Manchu collection of the St. Petersburg Institute of Oriental Studies.

The versions of some of them could be found in the other collections of the world. Their comparative study may enrich our knowledge of history and culture of the Manchus, as well as contribute to the sinological studies of the Qing China.

We limited ourselves to those which are registered in the "Description of Manchu manuscripts from the Institute of the Peoples of Asia" by M. P. Volkova. The present work with the collection showed a considerable number of manuscripts and block-prints which were not included into her both catalogues. Many of them deserve special future presentation.

Notes

1. The article deals with the material of the already printed Manchu catalogues of this collection, see M. P. Volkova, *Opisanie man'chzhurskikh rukopisei Instituta narodov Azii AN SSSR* (Description of Manchu Manuscripts from the Institute of the Peoples of Asia, USSR Academy of Sciences) (Moscow, 1965); M. P. Volkova, *Opisanie man'chzhurskikh ksilografvov Instituta vostokovedeniia AN SSSR* (Description of Manchu Block Prints from the Institute of Oriental Studies, USSR Academy of Sciences), 1 (Moscow, 1988). During the following studies it was discovered that these catalogues are not complete and miss many important texts. The work on the additional volume of the Manchu catalogue is in progress and the most interesting manuscripts of the new volume will be discussed in a separate article.

2. P. E. Skachkov, *Ocherki istorii russkogo kitaevedeniia* (Sketches on the History of Russian Sinology) (Moscow, 1977), p. 285; I. I. Zakharov, *Polnyi man'chzhursko-russkii slovar'* (A Complete Manchu-Russian Dictionary) (St. Petersburg, 1875), p. XXX; T. A. Pang, "A historical sketch of the study and teaching of the Manchu language in Russia (first part: up to 1920)", *Central Asiatic Journal*, XXXV, 1—2, pp. 123—37.

3. The document of N. N. Krotkov's release from the position of Russian consul in Urumci is dated by February 20, 1912, see Archives of Orientalists, file 32, opis' 1, No. 12. For the material referring to his diplomatic activity in Russian, Chinese and Manchu see Nos. 11—15. Many of N. N. Krotkov's letters and documents are at present in the City Archive of Krakow, Poland, as a part of V. V. Kotvich (1872—1944) personal archive.

4. *Protokoly istoriko-filologicheskogo otdeleniia Akademii nauk* (Protocols of the Historical and Philological Department of the Academy of sciences), September 16, 1909, paragraph 0238; *Protokoly russkogo komiteta dlia izucheniia Srednei i Vostochnoi Azii* (Protocols of the Russian Committee for the Studies of the Middle and Eastern Asia), May 10, 1908, No. 2, paragraph 28; September 22, paragraph 49 and others.

5. For the letters of S. F. Oldenburg to N. N. Krotkov see file 32, opis' 2, No. 9; for the letters of V. V. Radlov see file 32, opis' 2, No. 11.

6. *Pamiatniki indiijskoi pis'memnosti iz Central'noi Azii* (Indian Written Monuments from Central Asia), 2. Publication of the texts, research and commentary by G. M. Bongard-Levin and M. I. Vorobyova-Desyatovskaya (Moscow, 1985), p. 18.

7. About N. N. Krotkov see T. A. Pang, "Der Schamanenhof". Die sibemandschurische Handschrift *Saman küwaran-i bithe* aus der Sammlung N. Krotkov (Wiesbaden, 1992), pp. V—IX. — Shamanica Manchurica Collecta, Bd. 2; K. S. Iakhontov, "Svodnyi ocherk Rossiiskikh materialov po iazyku i kul'ture sibo" ("A union catalogue of Russian materials on language and culture of the Sibe), *Kunt-skamera. Etnograficheskie tetrady*, 2—3 (St. Petersburg, 1993), pp. 113—37.

8. File 32, opis' 1, No. 10.

9. M. P. Volkova, *Opisanie man'chzhurskikh rukopisei*; K. S. Iakhontov, "Svodnyi ocherk". Some texts not mentioned in the first catalogue and mentioned as lost ones in the second are discovered during recently done inventory. The description of these and other not registered Manchu manuscripts and block prints is under preparation.

10. The Oriental Institute in Vladivostok was established in 1899 as an educational centre for preparing students for administrative and commercial companies in Eastern Russia and the neighbouring countries. In 1920 the Oriental Institute was reorganized into the Oriental department of the Far-Eastern State University.

11. His scientific works are listed in *Bibliograficheskaia spravka, sostavlennaia obshchestvom Vostokovedeniia, dlia predstavleniia A. V. Grebenshchikova k rabote v AN SSSR* (Bibliographical note compiled by the Orientalistic Society to recommend A. V. Grebenshchikov for working in the USSR Academy of Sciences), file 75, opis' 2, No. 1.

12. For details see T. A. Pang, "Arkhivnye materialy A. V. Grebenshchikova po shamanstvu man'chzhurov" ("A. V. Grebenshchikov's archive materials on the Manchu shamanism"), *22 godichnaia nauchnaia konferentsiia "Obshchestvo i gosudarstvo v Kitae* (Moscow, 1990), part 3, pp. 136—40.

13. Grebenshchikov's translation is the only Russian translation of this text (file 75, opis' 1, Nos. 72, 73). The translations into French were made in the last century by L. Angles, *Rituel des Tartars Manchoux* (Paris, 1804), and by C. Harlez, *La religion nationale des tartars et orientaux mandchous et mongols avec le Rituel tartar* (Bruxelles, 1887).

14. M. P. Volkova, "Grebenshchikov Aleksandr Vasil'evich (1880—1941)", *Pis'mennye pamiatniki i problemy istorii i kul'tury narodov Vostoka. 19 nauchnaia sessiia LO IV AN SSSR*, part 1 (Moscow, 1986), pp. 28—9.

15. "Nisan samani bithe" — *Predanie o nishan'skoi shamanke* (The Epic Tale about Shamaness Nisan). Publication of the text, transliteration and Russian translation by M. P. Volkova (Moscow, 1961).

16. For literary discussion see St. W. Durrant, "The controversy among Western sinologists regarding the utility of Sino-Manchu translations", *Proceedings of International Ching Archives Symposium. July 2—6, 1978* (Taipei, 1978), pp. 66—80. The popularity of this text could be judged from the following publications: G. Stary, *Three Unedited Manuscripts of the Epic Tale "Nisan Saman-i bithe"*. Facsimile edition with translation and introduction (Wiesbaden, 1985); K. S. Iakhontov, *Kniga o shamanke Nisan'* (A Book about Shamaness Nisan) (St. Petersburg, 1992). — Folklor narodov Man'chzhurii, 1. For translations into different languages see G. Stary, "A new Altaistic science: "Nishanology"", *Altaica Osloensia. Proceedings from the 32d Meeting of the PIAC. Oslo, July 12—16, 1989* (Oslo, 1990), pp. 317—23. There are three Chinese translations of M. P. Volkova's text.

17. In Volkova's catalogue Sun Yun is mentioned as a Manchu. In fact he was a Mongol from Khorchin tribe which took the name Marat. Volkova's mistake was also noticed by Martin Gimm ("Zu den Mandjurischen Sammlungen der Sowjetunion, 1", *T'ong Pao*, LIV, 4—5, p. 295). For his biography see A. Hummel, *Eminent Chinese of the Q'ing Period*, ii (Taipei, 1991), pp. 691—2 (reprint edition); V. Veit, *Die Vier Qane von Qalqa*, Teil 1 (Wiesbaden, 1990), pp. 85—6. — Asiatische Forschungen, 111.
18. Sungyun, *Emu tanggu orin sakda-i gisun sarkiyān* ("The Tales of the 120 Old Men"). Introduction and facsimile by Kanda Nabuo (Taipei—San Francisco, 1982) (first publication); G. Stary, "Emu tanggu orin sakda-i gisun sarkiyān". Die Erzählungen der 120 Alten. Beiträge zur mandschurischen Kulturgeschichte (Wiesbaden, 1983). — Asiatische Forschungen, 83.
19. The texts of such kind were put down by F. V. Radlovskii, see S. Kaluzynski, *Die Sprache des Mandschurischen Stammes Sibe aus der Gegend von Kuldzha* (Warszawa, 1977). The texts of the Sibe stories and tales transliterated by V. V. Radlov are to be found in the Archives of Orientalists (file of A. O. Ivanovskii — file 20, opis' 1, No. 17) and in the Archives of the Russian Academy of Sciences (file of V. L. Kotwicz, 761). The V. V. Radlov's text of "Sidi Kur" copied by Kotwicz was published in facsimile with linguistic analysis, vocabulary and transliteration by E. P. Lebedeva and L. M. Gorelova, *Sidi kur: sibinskaia versia "Volshebnogo mertvetsa"*. *Teksty v zapisi V. V. Radlova* (Sidi Kur: A Sibe-Manchu Version of the Bewitched Corpse Cycle Transcribed by V. V. Radlov). With transliteration and English introduction by Giovanni Stary (Wiesbaden, 1994). — Aetas Manjurica, 4. For a review on this work see T. A. Pang, Review in *Manuscripta Orientalia*, No. 2.
20. *Kniga o shamanke Nishan'* (A Book about Shamaness Nisan). Facsimile, transliteration, Russian translation, commentary and foreword by K. S. Iakhontov (St. Petersburg, 1992). — Folklor narodov Man'chzhurii, 1; B. R. Ruchtsfeld, "Die Mandschu Erzählung 'Nisan saman-i bithe' bei den Hezhe", *Münchener Beiträge zu Volkskunde*, No. 2 (1989), pp. 117—55.
21. See also A. V. Grebenshchikov, "Kratkii ocherk obratzsov man'chzhurskoï literatury" ("A short sketch on the samples of the Manchu literature"), *Izvestiia Vostochnogo instituta*, XXXII, 2 (Vladivostok, 1909), pp. 19—20.
22. M. Gimm, *Die Literatur der Manjuren* (Wiesbaden, 1984), p. 209. — Neues Handbuch der Literaturwissenschaft, 23.
23. The facsimile of this manuscript with introduction, transliteration and German translation was published in Germany see T. A. Pang, "Der Schamanenhof".
24. Hewenjijun-Jalunga, *Saman jarin* (Urumci, 1990). For Chinese translation see Yun Zhixian, "Saman shengge", *Shijie zongjiao yanjiu*, No. 2 (1989), pp. 114—42; Qicheshan, "Sibo zu saman ge", *Sibo zu lishi yu wenhua* (Urumci, 1989), pp. 233—346. For German translation with commentaries see G. Stary, *Das Schamanenbuch der Sibo-Mandschurien* (Wiesbaden, 1992). — Shamanica Manchurica Collecta, Bd. 1.
25. For the first attempt to compare these texts see T. A. Pang, "New evidence about shaman practices among the Sibe-Manchu people in Xinjiang", *Proceedings of the 35th PIAC, September 12—17, Taipei, China*. Ed. Chieh-hsien Ch'en (Taipei, 1994), pp. 387—97.
26. For Chinese translation see Guang Singcai, "Xi qian zhi ge", *Li he*, No. 2 (1981), reprinted in *Xibo zu wenxue lishi lunwen zhi* (Urumci, 1981), pp. 147—50 and *Shige zhi* (Peking, 1983), pp. 393—413. In the Sibe-Manchu language: *Sibe i gurinjihe irgebun*, published by Heyere Sinciyān in 1985 in *Cabcal serkin*, No. 1995 (6.11.1985), p. 3; No. 1996 (9.11.1985), pp. 3—4; No. 1997 (13. 11. 1985), p. 3; No. 2001 (27.11.1985), p. 4; No. 2005 (11.12.1985), pp. 3—4. Guwan Singzai, *Sibe ukسرائ gurinjihe ucun* (Urumci, 1982), pp. 1—34.
27. G. Stary, *Epengesänge der Sibe-Mandschuren* (Wiesbaden, 1988), pp. 11—27.
28. In the article: Zhong Lu, "Sibo zu minjian wenxue jianjie", *Xinjian minjian wenxue*, No. 9 (1984), pp. 125—44; *Sibe ukسرائ irgen siden ucun* (Cabcal, 1984), pp. 84—9; *Kuren folon*, No. 1 (Urumci, 1983), pp. 99—134; *Sibe ukسرائ an tacin* (Urumci, 1989), pp. 279—89. Three versions of this text (not including our Institute texts) were translated by G. Stary, *Epengesängen*, pp. 73—130.
29. A. O. Ivanovskii, *Man'chzhurskaia khrestomatiia*, ii (St. Petersburg, 1895), pp. 191—2.
30. For transliteration and German translation of this text see G. Stary, "Der Mandschukhan Nurhaci als Held mandschurischer Lieder und Geschichte", *Fragen der mongolischen Heldendichtung. Teil IV*, ed. W. Heissig (Wiesbaden, 1987), pp. 180—205.
31. These Manchu songs were published in *Sibe ukسرائ irgen siren ucun* (Cabcal, 1984), pp. 30—3 and *Sibe ukسرائ an tacin* (Urumci, 1989), pp. 408—9. They do not repeat the Institute's text.
32. *Zhongguo shaoshu minzu wenxue*, vol. I (Changsha, 1983). For more details see T. A. Pang, "An introduction to the literature of the Sibe-People", *Central Asiatic Journal*, XXXVIII, 2 (1994), pp. 188—213.
33. T. A. Pang, The Manchu Manuscript "Manjugusai boo banjire nyiāman jafara dorolon i bithe" (The Book of Wedding Ceremonies of the Manchu Banner Families) (Wiesbaden, 1992), pp. 188—212. — Aetas Manjurica, 3.
34. T. A. Pang, "Eine unedierte mandschurische Handschrift über die Hochzeitbräucher der 'Mandschubanner' in Sinkiang", *Ethnohistorische Wege und Lehrjahre eines Philosophen. Festschrift für Lawrence Krader zum 75. Geburtstag* (Frankfurt/M, 1994), pp. 268—78.
35. S. Durrant, "Sino-Manchu translations at the Mukden court", *Journal of American Oriental Society*, No. 99 (1979), p. 653.
36. M. Gimm, "Manchu translations of Chinese novels and short stories: an attempt at an inventory", *Asia Major*, New Series, No. 2 (1989), pp. 77—114. There Prof. Gimm refers to the novel "Jeo pu tuwan-i bithe" as a lost one. Owing to the invitation of the Oriental faculty of the Warsaw University in December 1992, I worked with my Polish colleagues in the library of the Krakow University, where the personal library of the famous German manchurist Möllendorf is kept. Among his books we have found the copy of the mentioned novel, which was considered to be lost during World War II.
37. M. Gimm, "Manchu Translations", p. 82.
38. This kind of literary research is done in B. L. Riflin, V. I. Semanov, "Mongol'skie perevody starinnykh kitaiskikh romanov i povestei" ("Mongolian translations of the old Chinese novels and short stories"), *Literaturnye sviazi Mongolii* (Moscow, 1981), pp. 234—79; S. Durrant, "A note on translation of Chinese historical romance into Manchu and Mongolian", *Chinese literature: Essays, Articles, Reviews*, 1—2 (1979), pp. 654—6; W. Heissig, *Geschichte der mongolischen Literatur* (Wiesbaden, 1972); E. Haenisch, "Der chinesische Roman im mongolischen Schriftum", *Ural-Altaiische Jahrbücher*, No. 30 (1958), pp. 74—92.
39. L. Bese, "On the Mongolian and Manchu collections in the library of the Hungarian Academy of Sciences", *Jubilee volume of the Oriental collection 1951—1976* (Budapest, 1978), p. 59. The numbers mentioned there are No. 190, 194—202, 205 of 1930. When we checked this collection, we found additional Nos. 209 and 210 of 1930 and Nos. 305, 308 of 1931. All together 17 issues. For additional information see G. Stary, "Manju journals and newspapers. Some bibliographical notes", *Proceedings of the 28th PIAC. Venice 8—14 July, 1985* (Wiesbaden, 1989), pp. 217—32.
40. A. V. Grebenshchikov, *Man'chzhurskaia khrestomatiia* (Manchu Reader) (Vladivostok, 1913), pp. 119—36.

Illustrations

- Fig. 1.** *Emu tanggu orin sakda i gisun sarkiyān.* Manuscript department, call number B 15, fol. 1a.
- Fig. 2.** *Siyoo wen di juktehe-i kumun maksī be toktobuha wasimbuha selgeri hese.* Archives of the Orientalists. File 32, opis' 1, No. 47, fol. 1a.
- Fig. 3.** *Sula i bithe.* Manuscript department, call number A 156, p. 4.
- Fig. 4.** *Ba na i ucun.* Manuscript department, call number A 51, fol. 1a.
- Fig. 5.** *Daicing guruni kasigar ucuni bithe.* Manuscript department, call number A 16, fol. 1a.
- Fig. 6.** *Muduri mukdehe uculen.* Manuscript department, call number A 53, fol. 1a.
- Fig. 7.** *Ice donjin afaha.* No. 187 of 1925. Manuscript department, call number C 1.
-

ORIENTAL MANUSCRIPTS AND NEW INFORMATION TECHNOLOGIES

Correspondence Round Table: Arabic/Farsi OCR

The Arabic OCR was and still is one of the most beloved babies of the group of St. Petersburg scholars. The transformation of the pure academic project into the software product with good market possibilities took place in a very short period of time and under the heavy pressure of "post-perestrojka" realities in the ex-USSR, sometimes sad, sometimes tragic... Ex-Soviet academic science and scholarship, Cinderellas at the ball of "new Russians", had to find their own way for survival in the situation of sharp budget cuttings — in 1995 the budget of several academic institutions with the world-wide fame was 200 times (!) less than that of 1985. The case of Arabic OCR, conceived and produced in St. Petersburg, is just one example of the new approach to the existing problems. To my mind, the final success of the project was based on the "three whales": the scholarly background of the specialists from the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, the talent and diligence of the group of young Russian programmers previously connected with ex-Soviet high-tech military industry, and last but not least, the broadmindedness of young and talented General Manager of al-Alamiah, Cairo, Dr. Ashraf Zaki. The OCR project is a good example of possible cooperation between different Russian and Arabic institutions, cooperation which is replacing now the years of ideologically based relations between the USSR and the Arab world.

After the appearance of Al-Qari al-Ali on the market I received several letters from my friends and colleagues who are working in different research centres all over the world. The problem of Arabic/Farsi OCR was also intensively discussed by the ITISALAT subscribers. The Editorial Board of *Manuscripta Orientalia* proposed to publish some of the reviews. This was the birth hour of the idea of "Correspondence Round Table". Among its participants I have to mention:

Joseph Norment Bell, Professor of Arabic, Head of Section for Middle Eastern Languages and Cultures, University of Bergen, Bergen, Norway

Tim Buckwalter, Arabic Multimedia Developer and Project Manager Alpine Media (ICUE), Orem, Utah, USA

Khosrow M. Hassibi, Senior Staff Scientist, PI of Arabic OCR Project, Automatic Document Recognition Group, Mittek Systems, Inc., San Diego, CA, USA

Jan Hoogland, Assistant Professor of Arabic Language at Nijmegen University, the Netherlands.

Alexander Matveev, post graduate student of the Oriental Faculty of St. Petersburg State University and special student of Prof. Joseph Bell from University of Bergen

Paul Roochnik, Moderator, ITISALAT, Co-Founder and Secretary of the Association for Arabic Computing in North America; Arabic Computational Linguist at Applications Technology (AppTek) Inc., McLean, Virginia, USA

Petr Zemanek, Deputy Head, Institute of Ancient Near Eastern Studies, Charles University Prague, Czech Republic

as well as Dr. *Badr al-Badr*, Dr. *Sabri Kawash*, Dr. *Abdel-Hadi* and several others.

The programs mentioned and/or implied in the discussion are as follows (first three phone numbers are in Egypt, check out also the web page at <http://www.rdrop.com/~abdu/arabic.html>):

Co.: **Asset**

☎ 247-3916/248-3187

Prog.: Iqra' (OCR)

Price: L.E 4600.

Co.: **Sakhr**

☎ 274-9929

Prog.: Al-Qari' al-Ali

Price: L.E. 3900

Bilingual (Can OCR Arabic and OCR English).

Co.: **IBM**

☎ 349-2533

Prog.: Qari' Al-Nosoo

Price: \$1100

Runs under OS/2.

Co.: **CTA, S.A.**

☎ 34-3-4850410

Prog. TextPert 3.7 Arabic

Requirements: Apple Macintosh; all more recent models.

I bear all the responsibility for combining the messages of one and the same author once sent to ITISALAT. The information of personal character was omitted as well as the passages not connected with the OCR. The style and language of electronic messages was preserved without nearly any changes.

The St. Petersburg OCR project still has its academic continuation. Thanks to financial support of the Norwegian Centre for Computing in Humanities we are close to the realization of the idea to produce the specialized software for the automatic shape analysis and classification of handwritten symbols in early Qur'anic manuscripts. The preliminary name of the program is ENTRAP and those who are interested in the basic peculiarities of the approach can consult the paper of Mr. Nourlan Kondybaev *Maximum Entropy Approach in Automatic Classification of Symbolic Images* in the Proceedings of the 4th International Conference and Exhibition on Multi-Lingual Computing (Arabic and Roman Script. L., 1994, 2.4—2.5). Of course, the results of the work will find their reflection on the pages of our magazine in the new year.

I hope that the subscribers of *Manuscripta Orientalia* will receive this issue at the end of December, so let me take the opportunity to wish all our readers and, especially, all Arabic OCR users the 100% advance recognition of all possible problems to appear in the New Year and successful spell checking of all the mistakes we have to do in our life.

Efim Rezvan,
Convener of the
"Correspondence Round Table"

SAKHR BILINGUAL OCR (AL-QARI' AL-ALI). A USER'S INITIAL IMPRESSIONS

In this paper I would like to record some initial impressions from working with the Sakhr bilingual OCR system known as Al-Qari' al-Ali, to comment certain specific features of the program and to suggest a number of ways in which it may be improved.

It is perhaps appropriate to start with a few remarks on the origin of the product. It was first mentioned in 1990, when Dr. Efim Rezvan of the St. Petersburg Branch of the Oriental Institute of the Russian Academy of Sciences proposed the development of such a program in his report "Computer Methods in Qur'anic Studies" presented at the 2nd Conference and Exhibition on Bilingual Computing in Arabic and English in Cambridge. Originally the program was conceived as a powerful tool to facilitate the preparation of critical editions of Arabic sources by means of transferring large amounts of printed Arabic texts to computer files for subsequent processing. The immediate objective was the preparation by Valeriy V. Polosin of a critical edition of the famous "Fihrist" by Ibn al-Nadim. Dr. Rezvan considered this to be an excellent opportunity to develop and apply new techniques and software, and managed to interest a group of talented young programmers who had worked in the former Soviet high-tech military industry in the project. For a year Alexander Staryh, Mikhail Beregov, Alexander Popov and Fedor Bikov, in collaboration with Efim Rezvan, devoted nearly all their free

time to the development of the DOS prototype of the program, which was given the name MULTREC (Multi-Lingual Text Recognizer). The program was demonstrated in 1993 at the 3rd International Conference and Exhibition on Multi-lingual Computing held at Durham, where it aroused considerable interest, since it was virtually the only working program of its type. At this time the software company al-Alamiah became interested in the program, and subsequent to a visit to St. Petersburg by al-Alamiah's General Manager Dr. Ashraf Zaki, the preparation of a new Arabized version of the program was planned. The new version combined the achievements of the Russian programmers with important contributions made by specialists at al-Alamiah.

The first commercial version of Al-Qari' al-Ali was marketed in 1994. This product, although quite useful, has not yet become wide-spread, on the one hand because of its recent appearance and on the other because of its relatively high price and the powerful hardware it requires (a Pentium processor and a scanner with 600 dpi resolution are recommended). Hoping to introduce the product to my colleagues in Arabic studies who may not have had the opportunity to use it yet, I would like to report briefly on some characteristics of the program and how it may be applied.

Characteristics and area of use

Al-Qari' al-Ali works under the operating system "An-Nawafidh al-'Arabiya" 4.01 (or later), which, in turn, is installed over a Windows 3.1 operating system. It allows the transfer of scanned images of printed Arabic materials into text format, yielding 8-bit encoded text files which can be processed with al-Alamiah's word processor "al-Ustadh" or, for example, with the Arabic version of Microsoft Word for Windows 6.0. The program can be used for recognizing any Arabic printed matter. But if the text contains numerous ligatures, which is characteristic of older printed texts [1], errors at "Recognition" are practically inevitable, so the user has to correct them later during "Spell Checking". The best results are obtained from well printed mod-

ern texts with a minimum of ligatures. It is possible to transfer rather quickly a modern book or magazine into computer text with few errors (no more than 1%). As for poorly printed older books with a great many ligatures not included on the training keyboard and a variety forms for a given character, the process of recognition is regularly accompanied by errors. With such materials the production of a computer text file is extremely time consuming because of the need for careful correction of the recognized text (at first with the help of the built-in spell checker, and then by checking the corrected text in Word 6.0 or some other word processor). Even so, the production of an Arabic text is much faster than by typing, though it requires a

more highly qualified user. *The use of such a program is at any rate practically the only way for the majority of European Arabists to computerize a large amount of Arabic printed matter.* For the majority of Eastern European scholars, moreover, the services of professional Arabic typists are beyond reach, and the percentage of errors in typed text is rather high as well. It should be noted, however, that the advantage in speed becomes significant only when transferring rather considerable amounts of text (not less than ten or twenty pages), because preparing the pro-

gram to work, that is, "teaching" a new font, is a laborious process with fonts of any complexity.

The program is particularly important for the urgent task of compiling databases of medieval and modern Arabic texts, such as the database being developed at the University of Bergen under the direction of Prof. Joseph N. Bell, where I had the opportunity to work with Al-Qari' al-Ali [2]. Another particularly promising database project was begun in November 1994 in Saudi Arabia [3].

Hardware requirements

Anyone who has dealt with Arabic PC software knows how complex and slow these programs tend to be, especially in comparison with similar Latin programs. Al-Qari' al-Ali is no exception to this rule [4].

The program requires at least 386 processor with 4 Mb RAM and 10 Mb available disk space, but a more powerful hardware configuration is very welcome. Working with a Pentium 90 with 16 Mb RAM produces quite acceptable results. On a 486 DX2/66 with 12 Mb RAM the processing of scanned images of Arabic text was less successful. The teaching of a font and the further recognition of scanned text were not very difficult for the smaller computer, although the processor could not provide permanent support for the keyboard layout display on the screen and restored it after each operation. But the subsequent spell checking takes far too much time. Going from one error to the next takes up to half a minute, and if the font requires further teaching, which is practically inevitable even for rather

carefully taught fonts, the total correction process for one word can take several minutes. Thus the use of a Pentium with 16 Mb RAM is to be recommended when working with the program.

However, I would like to emphasize that a powerful processor is required not primarily for recognition of the text, but first of all for the spell checking, which is most important when working with poorly printed materials and complex fonts. In the case of modern books or typewritten texts, a weaker configuration (486 or even 386 with 4—8 Mb RAM) can be used.

The scanning resolution recommended by the manual is 300 dpi. However, it seems that the scanning of rather complex fonts of small size with this resolution can cause too many errors during the recognition process. In such cases, if the hardware configuration allows normal work with a higher resolution, this would be preferable. I achieved acceptable results scanning with 600 dpi.

Some remarks on the work with the program

Al-Qari' al-Ali comes with a standard set of modern computer fonts, which the program can recognize automatically. If the font of the scanned text is not included in this set, the program, after a search which may take some minutes, reports that no built-in font coincides with the scanned one. In such a case one must teach the program the new font. For this purpose it is generally sufficient to process in "learning" mode at least one and one-half to two pages of text, after which almost all letters, ligatures, and other symbols of the font will have been taught. Afterwards, it is useful to process one or two additional pages in a separate recognition mode within the learning option.

Learning option

In the learning mode each character or ligature of the scanned text is distinguished by the program, and the user must choose its alphabetic equivalent from the four-page keyboard layout on the screen (letters are on the first page, figures and other special characters on the second, ligatures on the third and fourth). At first one must do this for every character. Eventually the program will offer its own choices, which one can accept, if correct, or replace.

In the special recognition mode within the learning option, the program, having already been taught most of the characters, automatically recognizes them, stopping only on the symbols which it can not recognize. The user can then set the alphabetic equivalent of the unknown symbol himself (as in the learning mode). This option

makes it possible to process a page quickly and to teach the program most of the remaining symbols. One should not pass to this mode too early, however, because during "recognition" the program may make mistakes which it will not be possible to discover automatically. The most usual errors are connected with diacritics, the first and main parameter analyzed by the program being the "shape" of the letter or ligature. Thus, for example, if the program has been "taught" medial "bā" but not medial "nūn", "yā" and other similar characters, it will consider every "one-tooth" letter as "bā". The same applies to such pairs as "r" — "z", "d" — "dh", and "f" — "z". This problem is particularly troublesome with ligatures. If, for example, the user has taught the program the medial ligature "nb", he must theoretically teach it another thirty-five similar ligatures, that is, all thirty-six two-letter combinations of the six "one-tooth" letters "b", "t", "th", "n", "y" and "hamza". Otherwise errors such as *kunba* for *kunya* will occur regularly throughout recognition. But the number of non-standard ligatures one may teach the program is limited to about 130, which is often insufficient for a complex font. Therefore, the user should consider very carefully whether it is necessary to teach a given ligature or not. The criterion, naturally, is the existence and frequency of similar ligatures. For example, if one has taught the ligature "bah", "yah", etc., it will be expedient to add the ligature "ah", though it is not a frequent one. If one has taught a ligature and sacrificed its uncommon variants, error when these rare combinations occur will be inevitable.

To avoid such errors the user must periodically go out of recognition mode and return to learning in order to check how correctly and completely he has taught the program the font in question. If too many errors occur, it is reasonable to remove some of the most "dangerous" ligatures. Sometimes it is better to get a symbol for an unrecognized character ("^") instead of the wrong letter or ligature. It is easier to find this symbol after recognition is complete and to correct it then. Spell checking could pass over possible errors, because it is not rare in Arabic that a change of a letter produces a new "correct" word. However, the teaching of the majority of such ligatures will usually be finished during the following step of the work, the recognition of the whole text.

Recognition option

Having taught the program the font, one can pass to the next step, namely recognition of the text. After recognition of a given page one should spell check the recognized text. The parallel movement of the cursor, which highlights a block of the text in two windows (*Text* and *Image*) makes it possible to correct errors rather easily. While correcting one should continue teaching the font, since from the *Spell Checking* window it is possible to switch to learning mode and teach incorrectly recognized characters. Usually, such errors are caused by ligatures, so at this step one will face the serious problem of selecting ligatures to be removed, because the number of ligatures that can be taught in one font is limited.

Spell checking (including the *final stages of teaching a font*) takes from ten to fifteen minutes up to one and a half hours per page initially, depending on the complexity of the font. As one progresses in a text, this time is reduced as the number of characters that must be taught decreases. However, for complex fonts the process of re-teaching continues, practically as long as one is working with the text.

When the user passes from the *Spell Checking* window to the learning option, he can also correct errors noticed in

the *Text* window (unfortunately in this version of the program it is impossible to pass directly from the *Spell Checking* window to the *Text* window). Here it is necessary to be very cautious. Firstly, the position of the cursor on the screen after moving it in the *Text* window does not correspond to its real position (the difference is some three or four places), so in order to find out where the cursor really is, it is necessary to perform some operation in the *Text* window or simply to move the text in the window slightly. Secondly, it is not a good idea to correct text below the last place checked, because it could cause loss of connection between the *Text* and *Image* windows before the text is fully corrected.

Once spell checking is completed, it is necessary (using a special option) to detect unrecognized characters (designated by the symbol "^") and to correct them.

The result is a quite readable text with rather few errors. For a font of average complexity they will number from ten to twenty per page. For simple modern fonts the number will be very small (as in similar Latin OCR programs), but for complex fonts the amount of errors can be rather considerable. Subsequent manual correction of the text in Arabic Word 6.0 (or other word processor) is recommended in all cases.

Once the processing of a page is finished, it must be added to the text file, which should be saved in the OCR program as Arabic MS-DOS Code Page 720, and opened as the same in Word 6.0. The most convenient way is to use files with 7—10 pages, since the further processing of larger files (more than ten pages) in Word will be quite slow.

If one has to convert a rather short text (no more than ten pages), it would be useful to reduce the time of teaching the font and to correct inevitable errors by using the spell checking application. In such a case it will be enough to teach only one page and then to process another in the recognition mode within the learning option. Thereafter one can turn to the recognition of the whole text with a minimum teaching of the font during spell checking.

Recommendations for improvement of the program

Since the program will continually be improved, we would like to point out some problems which we hope the developers will take into consideration in future upgrades.

1. Switching From Spell Checking to Text Window

The most needed improvement would be to provide direct switching from the *Spell Checking* window to the *Text* window without closing the former (a similar function exists, for example, in Word 6.0). This is especially important because of certain peculiarities of the Arabic script. In a great many old Arabic printed texts the spaces between words are not indicated, and spaces often occur in the middle of words, rather than between them. As a result, the word could sometimes be cut in two. For example, the word 'arsala could be cut in two, if the break between "rā" and "sin" is too large. Such an error, although quite typical, will never be found by the spell checker, since both *sall* and *'ara* exist in Arabic [5]. The second half of the incorrectly divided word 'aw[[dahnā will be discovered by spell checking, since the word "d-h-n" does not exist in Arabic, but nevertheless one will not be able to correct this error from the *Spell Checking* window. The user has to

delete the space between 'aw and d-h-n, but for this purpose it is necessary (1) to close the *Spell Checking* window, (2) to pass to the *Text* window, and (3) to delete the space. Afterwards the user has to start spell checking again, so the operation will take considerable time. The fact that many scanned texts will contain a large number of such errors is the main reason why easy and fast switching to the *Text* window is desirable.

A second method to solve this problem would be to provide "Delete Space Back" and "Delete Space Forward" options inside the *Spell Checking* window.

2. Zooming of the image

In the present version of the program, zooming of the text image enables the user to enlarge it (in the learning mode and while spell checking), but does not allow him to diminish it. However, if only a very small part of the image is visible in the window, there is often a need to zoom the text out [6]. So, it would be useful to add to the zoom feature one or two options less than 100%, including at least one as low as 75%.

3. Learning option

a. When switching from learning mode to recognition (and vice versa), it is necessary first to stop the learning process, then switch to recognition mode, and then start the learning process again (altogether 3 steps). This switching operation takes a great deal of time, and in the second stage of teaching a font it must be carried out quite often. Direct switching from learning to recognition (and back) by pressing a key would be a considerable improvement.

b. A space symbol should be included in the keyboard character set. Sometimes it is senseless to provide meaning for a "symbol" distinguished by the program (for example, if it is a printing error, non-text mark, just a paleographical or paper defect, etc.).

c. A feature making it possible to "erase" such unnecessary elements from the scanned image would also be useful. An eraser is likewise necessary when adjusting the frame before the recognition of text (especially when dealing with poor quality printing, where the Arabic text frame is accompanied by a "dust" cloud), because the options "Selection of Text for Recognition" and "Marking Zones to Be Excluded from Recognition" are sometimes too cumbersome.

4. Ligatures

a. The window listing optional ligatures gives them in the order of creation rather than alphabetically, which in most cases makes it more difficult to find the ligatures one is looking for. Therefore, if possible, the ligature window should include an optional alphabetical sorting button.

b. In the current version, after a ligature window is opened and, then closed or removed, the cursor moves back to the top of the ligature list. In order to continue work with the ligatures, one has to place the cursor in the window of the ligature list and click it. Then one must start looking through the ligature list from the very top (this is especially important when the number of non-standard ligatures has already been exceeded and one has to select ligatures to be removed). It would be much more convenient if the cursor returned to the former position in the ligature list after closing or removing an opened ligature.

c. It would be helpful if the following standard ligatures, or at least some of them, were included in the third and fourth pages of the keyboard layout display (a fifth page would be a useful addition):

- 1) "ك + كم; كما; كلا; كل; كا; ... + ك":
- 2) "خى: حى (three variants — حى: هى: كى: ... + خى): عى: عى: مى: (two variants):
- 3) "به/بة": twelve variants (although the actual number of quite frequent ligatures is smaller) + عه (four variants); مه (two variants); له (two variants); كه (two variants);
- 4) "م + ...": عم (two variants); عم: مم: م: ... (two variants); هم (three variants); سم: هم (two variants):
- 5) "لح": three variants;
- 6) "ح + ...": صد (six variants); صد: جد (two variants); جد (six variants);

7) بنى: على: محمد: لله:

8) The dialogue boxes for some standard ligatures (for example, "b-y") offer only two standard positions (final and isolated), but in some fonts other positions occur, so these non-standard positions should be included as well. In the current version of the program the only way to deal with such positions is to create three-character non-standard ligatures, such as "b-y-keshida".

d. The fact that the program separates poorly the following ligatures should be considered:

- 1) "Consonant + alif" (for example, "nā", "fā", etc.)
 - 2) "Consonant + rā" (for example, "l-r", "m-r")
 - 3) "Consonant + wāw" (for example, "bū", "fū")
 - 4) "Consonant + sin" (for example, "yas", "fas")
 - 5) "Consonant + "single-tooth" consonant" (for example, "kan", "man")
 - 6) Two "single-tooth" consonants (for example, "yab").
- This is probably a function of the peculiarities of the fonts, on which I tried the program, but nevertheless it would be useful if the separation of such frequent quasi-ligatures were improved in the next version of the program.

e. It would be helpful to have two operating modes:

- 1) one mode with a minimum of non-standard ligatures, some fifty to seventy (for a simple modern font);
- 2) one mode with a maximum of non-standard ligatures, some 200 to 300, or more (for complex fonts).

The availability of two such operating modes would, on the one hand, simplify and accelerate using the program with modern texts. On the other hand, it would also facilitate work with complex fonts, because the amount of permitted non-standard ligatures (about 130) is not always enough, even for unvowelled texts, and it is obviously not sufficient for work with fully vowelled texts such as poetry.

Of course, an indefinite number of allowed non-standard ligatures would be welcome, since work with complex multi-ligature fonts would become much easier, though increasing the number of ligatures taught, if possible with the current algorithm, will slow the program down correspondingly.

f. There are also some problems related to the recognition of separate consonants:

1) The program unsatisfactorily distinguishes a medial "hā". Despite the presence of a specific variant of this glyph among the variants taught, the program often fails to recognize it. When this happens, increasing the number of variants taught provides little help.

2) The same problem occurs with "kāf" and its ligatures ("kā", "kāl", "kām" and others). In the case of "kāf", the problem is probably caused by the link of the "tail" of the "kāf" with the previous letter, but in the case of "hā" the reason is not clear. Perhaps this peculiarity of the program should simply be accepted and the user should not continue to teach these two letters if the program fails to learn them.

g. It would be useful to increase the number of variants allowed for two letters: "alif" (isolated and final) and "lām" (initial and medial). At least twice the normally permitted eleven variants are required, since confusion of these two letters is common in many fonts. In the current version of the program one must select very carefully which variants are to be kept or removed, a process which is time-consuming, but still produces only meagre results.

مادام بطناً وأذا بس تخامته وهو موم فأول وقيل هي شجرة ناز به تشبه الضربيع وقال ابن ٤٤٩ كيسان هو طعام بضربون عنده
 وبذون وبضربون إلى
 انه تعالى ملأه بالخلوص
 منه فسمى بذلك وهذا
 طعام لبعض أهل النار
 والرقوم والغائبين
 لا تحزين (لابس) ولا
 يس من شأنه الايمان
 والاشباع كما هو شأن
 طعام الدنيا وانما هو شيء
 بضربون إلى أكله من
 غير ان يسكون له دفع
 اخر ورتهم لكن لا على
 ان لهم استعدادا للشيء
 والسن الا انه لا يقدم
 شيئا مما يمل به الى انه
 لا استعداد من جهه
 طه لهم ويحقق ذلك
 ان جوهم وعطشهم
 لبسان قبيل ما هو
 المهم ومنه ما في هذه
 النشأة من حالة عارضة
 للانسان عند استعداء
 الطعم عند ما يقبل
 من البدن مشوقه الى
 المقهور المترتب بحيث
 يتلذذ به ما عند الاكل
 والشرب ويستفيح بها
 عند غيها عن سد
 استقرارها في العدة
 ويستند به ما قوتها
 عند غيها عن سد
 جوهم عبارة عن
 اضطرابها عند اضطراب
 التارفي استحاشها الى
 ادخال شيء كسيف يلوها
 ويخرج ما في باطن
 الهب واما ان يكون لهم
 شوق الى طعام ما وانما نذبه عند الاكل واستغناؤه عن غيرها واستعداد قوتهم

لا يمكن الذين كفروا من اهل الكتاب والمشركين من تفكيك حتى تأتيهم البينة رسول من الله يتلو صحفا
 مطهرة فيم اكتب فيها وما تفرق الذين اوتوا الكتاب الا من بعد ما حادتهم البينة مع اعلم ان في الآية
 مسائل (المتعلقة الاولى) قال الواحد في كتاب البسيط هذا الآية "من اصاب ما في القرآن نظاما
 ونفسه او قد تحفظ من الكفار من العلماء ثم انه رجع الله تعالى لم يلخص كيفية الاشكال فيها وانا نقول وجه
 الاشكال ان تغدير الآية لم يكن الذين كفروا من تفكيك حتى تأتيهم البينة التي هي الرسول ثم انه تعالى لم
 يذ كر انهم منة يكونون عن ما ذكروه مع قولهم اذا ارادوا الكفر الذي كانوا عليه فصاروا لا يتدبر لم يكن الذين
 كفروا من تفكيك عن كفرهم حتى تأتيهم البينة التي هي الرسول ثم ان كلمة حتى لانها انما هي ذم الآية
 تعني انهم صاروا من تفكيك عن كفرهم عند انان الرسول ثم قال بعد ذلك وما تفرق الذين اوتوا الكتاب
 الا من بعد ما حادتهم البينة وهذا يقتضي ان كفرهم قد ازداد عن عيسى الرسول عليه السلام غيبثا يحصل
 بين الآية الاولى والاية الثانية منة ماضية في الظاهر هذا منتهى الاشكال فيما اظن (والجواب) عنه من
 وجوه (أولها) واسم الوجه الذي نخصه صاحب الكشاف وهو ان الكفار من الفرقين اهل الكتاب
 وعبدة الاوثان كانوا يولون قبل مبعث محمد صلى الله عليه وسلم لا تفك عما نحن عليه من دينه ولا يتفك
 حتى يمشي النبي اوجه الذي هو مكتوب في التوراة والانجيل وهو جرحه السلام فكيف الله تعالى
 ما كانوا يولونه ثم قال وما تفرق الذين اوتوا الكتاب يعني انهم كانوا يبدون اجتماع الكافة والاتفاق على
 الحق اذ اذاهم الرسول ثم ما فرقه عن الحق ولا فرقه عن الكفر الا على الرسول ونظيره في الكلام بان
 يقول الفقهاء الفاسي لمن بعد لمست امتنع مما افاد من الاقوال القبيحة حتى يرضق الله الفتي فيما رزقه
 الله التي ازيد ما في قوله واعقله لم تكن منة كائن الفسق حتى تفسر ما عشت وأسلق في الفسق الا بعد
 الباري ذكرا كما كان يقوله فوجها (اما وما حصل هذا الجواب بربيع الى حرف واحد وهو ان قوله لم يكن
 الذين كفروا من تفكيك عن كفرهم حتى تأتيهم البينة مع كونه كونه عنهم وقوله وما تفرق الذين اوتوا
 الكتاب هو ان اربعين الواقعة والمعنى ان الذي وقع كان على خلاف ما دعوا (وثانها) ان تغدير الآية لم
 يكن الذين كفروا من تفكيك عن كفرهم وان جاءتهم البينة على هذا التقدير بزول الاشكال هكذا ذكره
 الفاضل الا ان تعديرا لفظه حتى هذا ليس من اللفظ في (وثانها) انما لا تحمل قوله من تفكيك على الكفر
 بل على كونهم من تفكيك عن كفرهم بما لما يقب وانه ضائل وما هي لم يكن الذين كفروا من تفكيك عن كفرهم
 محمدا بالمتاب والفضائل حتى تأتيهم البينة قال ابن عرفة أي حتى تأتيهم فاقطع لفظ المضارع وعناه
 الماضي وهو قوله تعالى ما تتلو التسابيح أي ما نالت والمعنى انهم ما كانوا من تفكيك عن كفرهم منة
 لما جاءهم بقره فواقبه وقال كل واحد فمعه قولا آخر يدنو بان نظيره قوله تعالى وكان من قبل يستمعون
 على الذين كفروا لما جاءهم من عرفة وكروايه والقول اختصار في هذه الآية هو الاول وفي الآية وجه
 رابع وهو انه تعالى حكم على الكفار انهم ما كانوا من تفكيك عن كفرهم الى وقت مجي الرسول وكيفية حتى
 تقتضي ان يكون الحال بعد ذلك بخلاف ما كان قبل ذلك والامر هكذا كان لان ذلك المجموع ما باء على
 الكفر بل تغرقوا فمهم من صاروا مؤمنين من صاروا كافرا وما لم يبق حال اولئك الجمع بعد مجي الرسول
 كما كان قبل مجيهم كفي ذلك في العمل بتدليل لفظ حتى ووقم اوجه خامس وهو ان الكفار كانوا قبل مبعث
 الرسول من تفكيك عن التردد في كفرهم بل كانوا جازمين به معتقدين حتى تزل ذلك الجزم بعد مبعث
 الرسول بل وقائما كمن تغدير من في ذلك الذين في سائر الاديان ونظيره قوله تان الناس آمنوا وحده
 فبعث الله النبيين مبشرين ومنذرين والمعنى ان الذين اوتوا الكتاب الذين كانوا عليه صار كما انه اختلط بهمهم ودهم
 طابودي كان ما في مود وكذا التصديقي وعاد الوثنيين لما بعث محمد عليه الصلاة والسلام اضطربت
 الخواطر والافكار وتشتك كل احد في دينه ودينه ومماتته وقوله تعالى من تفكيك مشهور به لان
 التفكيك الشيء من التي هو انفسه عنه فمات ان قلوبهم ما حدثت عن تلك الاقوال وما انفصلت

(٥٧ - غير نمران) شوق الى طعام ما وانما نذبه عند الاكل واستغناؤه عن غيرها واستعداد قوتهم

Fig. 1

(لم يكن الذين كفروا من أهل الكتاب والمشركين منفكين حتى تأتيهم البينة رسول من الله يتلو صحفا مطهرة فيها كتب قيمة وما تفرق الذين أوتوا الكتاب إلا من بعدما جاءتهم البينة) اعلم ان في الآية مسائل (المسئلة الاولى) قال الواحدي في كتاب البسيط هذه الآية من أصعب ما في القرآن نظما وتفسيرا وقد تخبط فيها الكبار من العلماء ثم انه رحمه الله تعالى لم يلخص كيفية الاشكال فيها وانا أمول وجه الاشكال أن تفد يرا لآية لبيكن الذين كفروا منفكين حتى تأتيهم البينة التي هي الرسول ثم انه تعالى لم يذكرنا نهم منفكون عن ما ذا لكنه معلوم اذ المراد هو الكفرا لذي كانوا عليه فصارا لتقدير لم يكن الذين كفروا منفكين عن كفرهم حتى تأتيهم البينة التي هي الرسول ثم ان كلمة حتى لا تنتها الغاية فهذه الآية تقتضي انهم عاروا منفكين عن كفرهم عند اني ان الرسول ثم قال بعد ذلك وما تفرق الذين أوتوا الكتاب إلا من عاروا ما جاءتهم البينة وهذا يقتضي أن كفرهم قد ازداد عند مجيء الرسول عليه السلام فحينئذ يحصل بين الآية الاولى والآية الثانية منافضة في الظاهر هذا منتهي الاشكال فيما أظن (والجواب) عنه من وجوه (أولها) وأسنها الوجه الذي لخصه صاحب الكشاف وهو ان الكفار من الفريقين أهل الكتاب وعبداء لاوه ثان كانوا يقولون قبل مبعث محمد صلي الله عليه وسلم لا ننكح عمار نحن عليه من ديننا ولا نتركه حتى يبعث النبي الموعود الذي هو مكتوب في التوراة والانجيل وهو محمد عليه السلام فحكى الله تعالى ما كانوا يقولونه ثم قال وما تفرق الذين أوتوا الكتاب يعني انهم كانوا يعدون اجتماع الكلمة ولا تفاق علي الحق اذا جاءهم الرسول ثم ما فرقهم عن الحق ولا أقرهم علي الكفر الا مجيء الرسول ونظيره في الكلام أن يقول ا لفقيرا لفاسق لمن يعظه لست أمتنع مما أنا فيه من الافعال القبيحة حتى يرزقني الله الغني فلما رزقه الله الغني ازداد فسقا فيقول واعظه لم تكن منفكا عن الفسق حتى توسر وما غمت رأسك في اللفق الا بعد اليسار يذكره ما كان يقوله توبيحا والزاما وحاصله ذا الجواب يرجع الي حرف واحد وهو في قوله لم يكن الذين كفروا منفكين كفرهم حتى تأتيهم البينة مذكور حكاية عيه م وقوله وتفرق الذين أوتوا الكتاب هو خبار عن الواقع والمعني ان الذي وقع كان علي خلاف ما ادعوا (وثانيتها) ان تفد يرا لآية لبيكن الذين كفروا مفكين عن كفرهم وانجا نهم البينة وهلي هذا لتقدير يزول الاشكال هكذا ذكره اللفافي الا أن تفسير لفظه حتى بهذلي من اللغة في شيء (وثانيتها)

Spell Checking

1. An on/off button for “*Suggestions*” would be useful (a similar option exists, for example, in Word 6.0). The function *Suggestions* is practically unnecessary, as the user can see the right word in the *Image* window, and it is moreover only of use when working with simple fonts. With complex fonts, the errors are so unpredictable that the user would hardly ever accept the suggestions proposed, although the program uses a great deal of time to produce them. However, spell checking with *Suggestions* in Word 6.0, when most recognition errors have already been corrected, is generally quite helpful.

2. The possibility to add new words to the spell checking custom dictionary would be of much use. A number of foreign borrowings, proper names, geographical names, and the like occur frequently, causing the spell checker to search again and again for the same unlisted item. With slower machines, the time spent can be up to ten or fifteen minutes per page.

3. Sometimes it is necessary to undo a correction or to see a corrected word again (as is possible, for example, with the “*Undo Last*” function in Word 6.0). The connection to the *Image* window in the case of the given correction will be lost, but the opportunity to go back would nevertheless be useful.

4. There are problems with the placement of the *Spell Checking* window. This window sometimes covers the text, and if one moves it, it will return to the former position after the next operation. The need to consult the context of a word to be corrected arises frequently. There are two possibilities here: (1) automatically moving the window to the top of the screen when spell checking reaches the middle of the page, or (2) saving the window position (i. e. not returning to the former position after the next operation). (It may be useful to create a button “*Save Spell Checking window position*”). The same problem exists, incidentally, with the *Find* and *Replace* windows.

5. It would be helpful to solve the problem mentioned above of the improper position of the cursor in the *Text* window, which occurs when the user during spell checking switches from the *Learning* window to the *Text* window.

My last recommendation concerns the program as a whole and the very principle of the recognition of Arabic symbols. Perhaps the technical implications are too great and would cause a considerable slowing down of the program, but it would be helpful if the program could take into account the position of a letter in a word (or a “block”) more precisely, considering the previous letter as well as the subsequent one. What I am suggesting is that not only the “shape” of a glyph should be taken into consideration, but its “position” in the word as well. Many errors could be avoided if position was taken into account.

For example, in a number of fonts in poorly printed texts, there is practically no difference between the shape of medial “‘ayn” and that of final or isolated “hā””. The reader can only understand the meaning of such a symbol according to its position in the word [7]. A similar analysis should presumably be done by the program. If a symbol being analyzed is followed by a medial or final variant of a letter, it means that the symbol cannot be final or isolated “hā””, but only medial “‘ayn”, even if the program can detect no difference in their shape.

However, such an analysis will not only require that the already recognized previous symbol be taken into account, but that the not yet recognized subsequent one be considered as well. The analysis of one symbol would thus consist of at least three additional steps (analysis of the letters on either side and of the group together). Since this would further complicate the program and would require even more powerful hardware, it is unclear whether such an innovation is feasible at the present time.

Notes

1. For example, works printed in the late nineteenth century in relatively complex fonts such as al-Razi’s *Tafsir* printed in Cairo in 1308/1890—1891, on which I tried the program (see *fig. 1*).

2. I would like to use this opportunity to express my gratitude to the Research Council of Norway, to the University of Bergen, and personally to Prof. Joseph N. Bell for giving me a chance to participate in this very interesting project.

3. The Research Institute for Computer and Electronics (RICE) at King Abdulaziz City for Science and Technology started in November 1994 a project to compile a large database of Arabic texts of different types (classical, modern, scientific, etc.) which will be available to all researchers doing Natural Language Processing research.

4. A common Latin OCR program can rather easily disassemble a word into vertical segments (characters), as the characters are separated by blanks, but for Arabic text this is an extremely sophisticated problem. Absence of blanks between the characters, overlapping of two (or more) characters or the parts thereof in one vertical segment (for example, a “tail” of a “kāf” and a previous letter), a multitude of diacritical marks, variant forms of the same letter, standard and non-standard ligatures, and so on make it necessary to compute many parameters at once and require a powerful processor and considerable RAM.

5. Another characteristic example of this kind can be seen in the accompanying “recognized” page from al-Razi’s *Tafsir*, where ‘alayhi as-salam has turned in ‘aly h as-sa lam” (see *fig. 2*).

6. If the user scans with 300 dpi resolution, this is not so important, but when he works with 600 dpi resolution it becomes necessary.

7. And, of course, from the context, but unfortunately we cannot use this criterion in the program.

Illustrations

Fig. 1. Page from al-Razi’s *Tafsir* (Cairo, 1308/1890—1891).

Fig. 2. The same page as “recognized” and spelled by Al-Qari’ al-Ali.

TEST OF TWO ARABIC OCR PROGRAMS

From the fourteenth to the eighteenth of December we met in Bergen to experiment with the two of the OCR programs for Arabic that were available in the software market as of November 1994. One of these was TextPert 3.7 Arabic, produced by CTA, Inc., which runs on the Macintosh Arabic system (system 7.1 was used in the test) [1]. The other was al-Qari' al-Ali (Arabic "Automatic Reader") 1.0, upgraded to 1.1, a version of the program known as MULTREC. It is produced by al-Alamiah Software Co. and runs on al-Nawafidh al-Arabiya, the Arabization program for Windows from the same company [2]. Taking part with us were administrative assistant and librarian Awni Taki Musa and undergraduate student Navid Saminasab.

The limited time and means at our disposal did not allow us to try out a third program, ICRA 4.0, which is an application for Windows (with Arabic Support) produced by Arab Scientific Software & Engineering Technologies (cf. the communications by Jan Hoogland, Discussion Fo-

rum on Personal Computers Arabization, Dec. 21, 1994; Itisalat, Jan. 5, 1995). Subsequently Jan Hoogland was himself able to compare al-Qari' al-Ali with ICRA, and found al-Qari' al-Ali to be superior (1) in character recognition, (2) in training for ligatures, (3) in the fact that the spelling checker is linked (initially) to the original image, and (4) in that the batch mode does not require confirmation after every page (cf. Itisalat, May 4, 1995).

Another program which has been discussed recently, one using neural-net based software from Mitek Systems in San Diego, was as of late November not yet available, and the company could provide no comparison results.

Both of the programs we tested were able to recognize certain computer printed texts of good quality with a reasonable degree of accuracy considering the difficulties of the Arabic script. Both were many times slower than comparably priced programs for Latin OCR, also when reading Latin.

TextPert

TextPert is a program which is extremely easy to use, but which offers in the normal version no means of influencing character recognition other than adjustment of resolution, brightness, and contrast on the scanner. Thus it was not possible to choose, or to train for, the fonts we were scanning. On very good and simple texts the results were approaching acceptable standards, but on more complicated fonts the program recognized virtually nothing. Moreover, on the computers we used (a PowerBook 180 with 14 Mb of memory and an LC III with 8 Mb of memory), the program was not always able to follow the paths between the automatically established zones on the

document to be read. When it could not do this, the Macintosh would crash. There is a much faster and three or four times more expensive version of Arabic TextPert which uses a RISC board. We have been told by the company that it does not perform essentially differently from the cheaper version except for speed, but that they may allow access to the engine for certain purposes the user may require. For Macintosh users who only want to scan certain kinds of computer produced documents, TextPert may offer something approaching an acceptable solution, but it is to be hoped that future versions will take into account the need to train for different fonts.

Al-Qari' al-Ali

This program is based on a very powerful algorithm which seems to combine vector and bit-map analysis. In its first upgraded version it offers a number of means, although still not quite enough, of controlling recognition performance. Thus it is possible to select desired level of accuracy and to train for the majority of fonts in Arabic and in most other scripts. The results of an OCR operation can be controlled with a spelling checker that, while far from what one might hope for, is surprisingly good, par-

ticularly for controlling words that have run together. To facilitate comparison between the original scanned image and the text document, the spelling checker highlights problem areas simultaneously in both.

The texts on which we tried al-Qari' al-Ali were for the most part photocopies from works printed in the late nineteenth century in relatively complex fonts (for example Shaykh'zadah's *Hashiyah* on al-Baydawi printed in Constantinople in 1306/1888—1889). There were quite a few

breaks between letters, and spaces as often as not occurred in the middle of words, rather than between them. The results were none the less impressive, although anyone interested in scanning texts of this type must be prepared to invest a great deal of time both before and after scanning.

The text documents we produced using al-Qari' al-Ali were later converted for the Macintosh using a conversion table we made in Paradigma 2.0, a program designed by Espen Aarseth at the University of Bergen. The PC Arabic system handles ligatures and initial and final forms differently from the Macintosh, and word boundaries in the text document produced by al-Qari' al-Ali were often clear on the PC even when there was no actual space between the words. These boundaries disappeared when the text was converted for the Macintosh. Since at this stage in the program's development the adding and subtracting of spaces has to be done manually, it is probably better to carry out this part of the correction process on a PC, even if one intends to continue working on a Macintosh later on. We understand that a Macintosh version of the program is under development, but we have no information about how this particular problem will be handled. Perhaps the best results can be achieved, once preliminary spellchecking has been carried out, by converting the text from the Alamah Nawafidh Windows code page to that used by Arabic Windows 3.1 or 3.11 by means of the utility al-Muhawwil that comes with this version of Arabic Windows, and then continuing correction of the text in the Arabic version of Microsoft Word for Windows 6.0.

The very considerable amount of time it takes to train for new fonts, especially hand set fonts with many ligatures, is one of the main problems with al-Qari' al-Ali. Even when teaching Latin fonts the process went slower and the operations were more cumbersome than, for example, in the bit-map program ProLector, which, however, is considerably more expensive. Quicker routines for training fonts would be a great improvement. A feature that al-Qari' al-Ali has which is not in ProLector, is the possibility of editing bit-map models within the program and inserting them into a set of previously trained models. Al-Qari' al-Ali has an English menu option, so it can in fact easily be used by persons unfamiliar with Arabic. The Arabic menus in al-Alamah's Nawafidh Windows constitute only a slight problem for experienced Windows users who do not know the language.

Because the program, although slow, seems so powerful and so promising, we would like to note some problems which we hope the developers will take into account in future upgrades.

✓ **Manual.** Although the manual may look nice, it contains only very superficial information and needs to be entirely redone. An English version would also be helpful.

✓ **Image rotation.** We did not find, within the program, a tool for gradual rotation of the images to be scanned or read. Such a feature would make it easier to maintain a constant alignment of scanned images so that the program always sees the characters it is to learn or read from the same angle.

✓ **Recognition blocks.** Al-Qari' al-Ali places groups of connected letters into a green frame and what it thinks are individual letters within the group between horizontally adjustable red lines inside the green frame. Neither the width nor the height of the green frame can be manually

adjusted, which means that characteristic elements of a block are on occasion excluded or extraneous information included. Within the green frame, the program lets one know what it is taking as characteristic of a letter by outlining it in blue. It would be helpful, if it is possible, to have a means, in addition to the red lines, of activating or deactivating the blue outline where the program has made a mistake. The program will have certain difficulties with complex fonts until these problems are remedied. For the moment the best guideline seems to be not to override the program's choice when training any more than necessary, since it is not unlikely that it will make the same choice again anyway. When the program has seen a medial letter or ligature as one in isolation because of breaks in the word, for example, the "in isolation" choice at times has to be accepted. Otherwise the program may fail to read the letter or ligature, or read it as something else.

✓ **Fonts.** The program comes with few pre-trained fonts, and those it does provide are computer fonts with few ligatures. Given the amount of time needed to train fonts, the library of pre-trained fonts, particularly non-computer fonts, needs to be greatly expanded. Further, the program lacks an efficient means of visually comparing fonts to be read with the pre-trained fonts, since the font display window in the "create/emend" font library dialogue box gives an inadequate image of small fonts. Lastly, there is in the present version no means of scaling up or down previously trained fonts, which means that every font in every size has to be trained separately. However we have been told by the company that in the next version it will be possible to reproduce fonts in other sizes (plus or minus 2 points in either direction).

✓ **Confusing messages.** One problem we experienced with al-Qari' al-Ali was that when all the places allowed for the variants of a character in a given position had been used up, the warning that appeared was not always the same. A character may have eleven variants in each position (initial, medial, final, or in isolation). When we tried to teach a twelfth variant, the message occasionally stated that we had exceeded some other limit. The problem may have been insufficient memory in the computer we were using, or it may be in the program. In any event, when using the current version of al-Qari' al-Ali one should be aware of the possibility of inappropriate messages appearing.

✓ **Ligature dialogue boxes.** The dialogue boxes for certain combinations of letters, such as "fii", offer only the normal position option, in this case "in isolation" or "final", when in fact in some fonts other positions occur. An "other" button is needed here to allow for the less common options.

✓ **Ligature list.** The window listing optional ligatures gives them in order of creation rather than alphabetically, which in most cases makes it more difficult to find the ligatures one is after. The current method, however, makes it easier to correct mistakes one has just made. If possible, the ligature window should include an optional alphabetical sorting button.

✓ **Limited number of ligature possibilities.** The possibilities for creating (coding for) new ligatures are far too limited for fonts of any complication. It is possible to train new fonts just for ligatures, but this is a complex and

laborious process that inevitably involves duplication, unless a separate record is kept for all ligatures. We need up to 1000 more possibilities just for fairly complicated older fonts. If the possibilities were almost infinite, the program could be used to recognize an almost infinite number of signs, images, and symbols, and al-Alamiah would, as a result, have an almost infinite market for their product.

✓ **Space markers.** Because of the problem with spaces between and within words alluded to above, the al-Muharrir word processor that comes with al-Qari' al-Ali should have an option for marking spaces between groups of letters.

✓ **Stability.** The stability of the program, especially when communicating with the scanner, seems to need improvement. The problem may have been in Arabic Windows or in our hardware. When writing the first version of this review we were using a modest Olivetti 486SX/25 Mhz with 8 Mb RAM and a Hewlett-Packard ScanJet IICx. Subsequently we have used a Compaq XL 590 with a Pentium 90 processor and 16 Mb RAM. This

machine may be a little too fast for the program. Recognition speed has increased by a factor of about nine, but clicking with the mouse does not always stop the recognition feature in the training mode as it should.

We tested incidentally some of al-Alamiah's other software, in particular the word processor Al Ostaz, the Koran database for Arabic Windows, and the hadith databases for Arabic DOS. All of these were impressive products which should receive a warm welcome in any milieu, academic or religious, with a special interest in the Arabic and Islamic heritage.

This review was first made available on the Internet on the lists Reader (14.01.95) and Itisalat (17.01.95), and the original version is preserved in electronic form and in hard copy in the Archive of Electronic Publications of the Section for Middle Eastern Languages and Cultures, University of Bergen (<http://www.hf-fak.uib.no/institutter/midtspraak/aep.htm>).

Notes

1. Al-Qari' al-Ali 1.1

Producer:

*al Alamiah Software Company
al Alamiah Building
Freezone, Nasser city
Cairo, Egypt.
☎ 20-2-2749929
fax: 20-2-2740044
e-mail: alamiah@intouch.com*

Requirements:

- IBM compatible; minimum 386 processor; 486 or Pentium is to be recommended
- 4 Mb RAM, 8 Mb recommended, 16 Mb still better
- mouse
- Windows 3.1
- al-Nawifidh al-Arabiya Arabic interface for Windows, version 4.01 or later (same producer)

2. TextPert 3.7 Arabic

Producer:

*CTA, S.A.
c/Joan d'Austria, 68
08005 Barcelona
Spain.
☎ 34-3-4850410
fax: 34-3-4855327
e-mail: textpert.int@applelink.apple.com*

Requirements:

- Apple Macintosh; all more recent models
- 1 Mb RAM minimum, 2 Mb or more recommended
- at least 2.5 Mb on HD
- Arabic system 6.1 or higher, including 7.1
- hardware protection unit

ITISALAT OCR DISCUSSION

Announcing the Internet forum on Arabic computing, called: ITISALAT. ITISALAT, which made the pages of *Al-Hayat* (the world's most widely circulated Arabic language news daily) on 2 May 1995, promotes contact and stimulates the exchange of information in the field of Arabic computing. ITISALAT subscribers discuss the following topics and others:

- Arabic optical character recognition (Arabic OCR);
- Arabic computational linguistics;
- Arabic machine translation;
- Arabic computer-based training/education;
- Arabic software development;
- Computing developments in the Arab World;
- Arabic corpora compilation;
- Arabic data storage & retrieval;
- Arabic hypertext;
- Arabic code standardization;
- Arabic email.

The name ITISALAT is actually a play on words. On the one hand it derives from the Arabic word "ittisalaat" which means "connections" or "communications". But it also forms the acronym for *IT IS Arabic Language And Technology*. More than half the 225+ ITISALAT subscribers live in North America. The rest are spread throughout Europe and the Middle East, including Egypt, Iran, Israel, Kuwait, Morocco, Saudi Arabia, Tunisia, and the UAE. And although they hail from widely diverging professions,

such as computer engineering, Arabic linguistics, and library science, they do share a common objective: to further the development of Arabic computing. ITISALAT began operating in May 1993.

The same wave of computer network communication that has captured the American imagination has begun to sweep over the Middle East. More and more Arab countries have connected to the Internet in recent years, and it won't be long before Arabic email becomes the medium of choice for communication with our friends and colleagues from Morocco in the West all the way to Iraq in the East. We hope that ITISALAT can play a small part in the Arabization of the Internet. Time will tell. ITISALAT frequently operates in cooperation with The Association for Arabic Computing in North America, ICEMCO, and The International Association for Arabic Computing. It has no official links with these or other organizations, but we all complement each other, and share a similar goal: to further the development of Arabic computing.

Arabic optical character recognition constitutes one of the most important topics of discussion on ITISALAT. On the surface, the discussion has concentrated on comparing and contrasting the various Arabic OCR software available on the market, along the dimensions of price, platform, ease of use, and quality of results. I hope that the readers of *Manuscripta Orientalia* will find the summary of ITISALAT discussion interesting*.



Even though I may seem an expert on Arabic OCR, I must stress I'm just an end-user with experience with two OCR-programs. I've reached acceptable results by using Al-Qari, but still sometimes (like yesterday night) I curse the product. I haven't published on the subject yet, but I'm now in the stage of preparing a paper about this subject for ICEMCO. This is also going to be a practical report on working with Al-Qari (and some remarks about the other program Iqra'). I will present some results during the presentation of the paper...

...I think I may say I am an experienced user of Al-Qari' al-Ali (Automatic Reader) of al-Alamiya (Sakhr), and also of Iqra' 4. Al-Qari is the newest and best performing of the two, so since about 6 months I stopped using Iqra'. I realized that in my enthusiasm when I reported on 99% success rate in Al Qari al-Ali, I maybe forgot to mention the spell-checking as being part of the process. So to repeat: I copy a page, remove all 'noise', scan it to make a .pcx file. Then I have Al-Qari recognize the .pcx-file. This is done at an average speed of 80—90 characters per second. When the text-file is on the screen, I run the spell check utility, that has a relation with the picture-file. When the spell

* To subscribe, send the command (A) to address (B): (A) — subscribe ITISALAT your-1st-name your-last-name (B) — listserv@listserv.georgetown.edu. For more information, contact: Dr. Paul Roochnik, Moderator, ITISALAT email: roochnik@ios.com. Paul Roochnik.

checker runs into an unknown Arabic word, this word is both highlighted in the text-file (right half of the screen) and in the (enlarged) picture of the original (left side of the screen). In this way it is very easy to correct the misrecognized word. This may take some time though. I think an average of 8 pages an hour can be done in this way. But this work can be done by a young/cheap assistant, so that we researchers can lean back and think, or surf on the WWW. But if a misrecognized word is still an existing word, spell checker just proceeds assuming nothing went wrong. So 100% is not possible. I checked Abdel-Hadi's Home Page (for the OCR-report). Joseph Bell who reported on Al-Qari earlier this year is on this list too. I think he was less positive because the type of texts he is dealing with is completely different: old printing with all the disconnected characters etc. But Joseph also once wrote on this list that Al-Qari's capacity for storing ligatures is limited. I must say I did not run into this problem. First of all there is a number of pre-defined ligatures (all the usual ones, some 30 or 40 of them), but there is also the possibility of user-defined ligatures. I did not count them, but I think my most trained typeface contains at least another 50 or 60 ligatures (sometimes just combinations of 2 characters that are not ligatures at all, but the program always picks them together).

I admit this 99% (which is just an estimation) can only be achieved under most favorable conditions, i. e.: very good paper quality, very clear print, after copying the original and removing all noise (headers, footers, illustrations etc.) and learning the type-face for several hours (especially when the typeface contains a considerable number of ligatures). The number I mentioned was realized on the Kuwaiti magazine 'Al-Arabi' which reaches us through the Kuwaiti embassy.

Jan Hoogland



Now a brief comment on the Arabic OCR programs: I have followed the reports on some of the OCR packages available for Arabic and posted on this list or some other similar lists but never felt the need to make any comments because I did not have any access to such programs for the simple reason that I could afford none of them. I am a user of OCR programs for Western languages and recognize their importance to anybody needing to move printed text to computer media.

My observation to Jan Hoogland is that his description of the product he reports on — with much appreciation from all of us — does not match high enthusiasm for the product. According to Jan Hoogland, it takes several steps of preparation of the text before it can be recognized by the OCR program, and the output of the program is about 8 pages per hour. Even after ...assuming that those pages are average printed book pages, I find the OCR program to be very slow and cumbersome. I managed better than 20 pages per hour scanning the text of Yusuf Ali's "The Holy Qur'an" including the foot notes. There was no need on my part to make any preparation of the pages which consist of three main parts each: English text, footnotes, and Arabic text. Many a time I scanned 60 or more pages at a time and unless I asked to defer recognition, the program would recognize all what I have scanned as soon as I stop scanning. The 20 pages per hour, I mentioned earlier from start to finish including the final correction and cleaning of the final text. At the end of all this I have the opportunity to save my work in any one of several formats including plain text.

From what I have read about Arabic OCR packages, there is nothing that may come close to what I have described, thus my question to the members of this list. Why should Arabic users of computers allow themselves to accept inferior products at exorbitant prices? Let us not blame it on the vendors. It is our responsibility to demand better products and refuse to use inferior ones. I would love to have a good and reasonably priced Arabic OCR program, but at the rate things are progressing in Arabic computing, I have a long wait ahead of me.

Sabri Kawash



I have been the PI of ARPA Arabic OCR project which started in July 1992 and ended in July 1995. The Arabic OCR research and the obtained results by other researchers by that time were poor. Most of them had used a very small database which had no real-world importance. My goal was to implement a system using neural network technology that can recognize average to good quality machine-printed Arabic text without retraining (the ultimate goal is a character accuracy of 98%). We had access to a database of 350 documents from magazines, newspapers, and books. The data was later enhanced by inclusion of 15 Windows fonts as well. The first Beta version of the system was made available in July 1995. ...As a researcher & a technology developer who has spent some time to develop algorithms for OCR of Arabic text, I have a few comments to make here. From a user point of view, your comments are true. The Arabic OCR systems are expensive and do not compare in performance to Latin OCR systems. But is this a fair comparison? — You have to consider that Latin OCR systems have been the results of hundreds of man years of research and development in at least last 30 years and millions of dollars in investment. Arabic OCR is in its infancy and there is not a whole lot of money being spent on it, — Latin text is much less complex to recognize compared to Arabic text or Chinese for instance. The complexity of Arabic/Farsi OCR is much more than Latin text OCR. You cannot compare apples and oranges. Also when the quality of Latin text documents decreases, the performance of these systems drops sharply as well, — the extent that a technology can be developed and enhanced is directly related to the investment in that technology. If there is money, it will absorb great talents to de-

velop and enhance the technology. To invest in any technology, one needs a reasonable rate of return on his/her investment. I do not know if the Arabic OCR market (low end users) is big enough to justify the risk for anyone. As a result, prices are high and performance is lower than comparable Latin OCR systems (add to that the non-existence of Software Copy Right Law). Still with those high prices, I am sure the vendors are not making money on it, — I have developed an OCR system for Arabic which is based on neural networks and does not require training — not sensitive to size or font- as long as the text is from Naskh family of fonts or something in that neighborhood (for other styles, a retraining is required). This is a superior technology compared to other Arabic OCR systems and is in Beta test now. Will it be available for end users any time soon given the market size? I guess not. Even Microsoft is not yet making any money on their Arabic products (correct me if I am wrong). Just imagine anybody who runs Arabic Windows is probably a potential Word for Arabic customer. When it gets to Arabic OCR, there is a small portion of this population that are potential OCR customers and currently it is not enough to make developers such as myself and technology investors excited. The high-end market, may be another story. In conclusion, I think if there was a Perfect Arabic OCR system, it would not make enough sales (as a low-end product) to support and maintain itself. So you are right that there is a long wait.... and this was our side of the story!

Khosrow M. Hassibi



Dear Dr. Hassibi,

Thank you very much for your reply and clarifications, but let me assure you that aside from your involvement in OCR software development, that all the points you made are well known to me and that is why I am so concerned about the cost and quality of software available for computer users who use the Arabic character set and its variations. The market is not insignificant, but not mature yet. Any investor in that market has to look for the future and not at the present. I don't know if Microsoft is making or losing money in that market, but they have managed to dominate it and set the standards for it while the users of those languages that use the Arabic character set are passively watching. Consider the mess we have with the serious standards created by serious vendors including Microsoft, and consider the disregard Microsoft exhibited when it disregarded the standards after they were set by introducing yet another standard for Arabic Windows. What objection was there from the users? None.

My criticism of the quality and the cost of software for the Arabic computer user — by Arabic I don't mean the language but the character set — was not aimed at the software developer or the vendor as much as it was aimed at the user.

My friends, technology is very expensive and cannot be developed at the expense of the end user especially during the development stages. Among the users of the Arabic characters there are many wealth countries, some of the wealthiest in the world. Certainly some of the governments of those countries can afford to make an investment in the future of their people and spend some money to develop some software that is commercially not feasible but important and necessary such as good OCR programs. Among the readers of this form there are people who are associated with some institutions in countries where millions of dollars are spent daily on less important things than the development of good software, and I am sure that some of those people are in a position to make such suggestions to their institutions and governments. But first a realization has to be made. If people do not demand quality for a fair price, they will continue to receive mediocre products at the exorbitant prices and that not everything can be bought at the open market. Something has to be home-grown and developed at home to meet the local needs.

Sabri Kawash



I think you made the point very clear. I 100% agree with you about what you have mentioned above. Guess some of these governments must invest in some of these technologies the same way that U.S. government has invested in areas of technology that did not yet have a mature market. Specially if these technologies are directly related to their culture and language. I hope your message will be heard.... Thank You for your thoughtful comments,

Khosrow Hassibi



The logical answer to this OCR debate is competition. If I were to develop a good Arabic OCR program, I would sell it for like \$200. The other OCR developers who are selling theirs for \$2000 will be forced, whether they like or not, to lower the price of theirs. Look around you and see how software and hardware prices have dropped dramatically because of fierce competition. Companies think of quality and customer service. If you're thinking of return on your investment, would you rather sell a \$2000 product to 1000 users or a \$200 product 10.000 users? When you're planning to develop an upgrade, which market would you choose, the 1000 users or 10.000 users? Once more good players join the game, the game will be more exciting!

Abdel-Hadi



Well. The point I was trying to make was that the market is small TODAY and it will be hard to attract new companies to COMPETE in it. TOMORROW things may be different. With such a small market, you will not get COMPETITION in near future.

Khosrow Hassibi



I don't think the market is that small that you can't justify the development of an OCR program. If you're reasonably priced, people will buy the product and you'll have a market. Some companies take the risk of putting out a new product where they don't know how successful it will be.

Abdel-Hadi



I do understand the remarks of both colleagues, but fully disagree on the point that OCR-software is of poor quality. Al-Qari' al-Ali performs very good and certainly meets my demands (although not all of them). And I do not think my demands are too low, what I need is machine-readable text that can be used for (manual) dictionary making. As I promised, I will keep you informed, and try to get some ftp-experience in order to make some files available in order for you all to judge yourselves.

Jan Hoogland



My research has focused on Arabic spell-checking, which I believe is an essential supporting component of successful Arabic OCR. However, effective Arabic spell-checking requires syntactic as well as morphological analysis, and I am unaware of any OCR system that incorporates both morphological and syntactical analysis. ...I have been impressed by Jan Hoogland's report for high OCR accuracy using Al-Qari' al-Ali (well, maybe not 99% accurate, but darn good), but the main problem for me is that I cannot see adding yet another OS (Sakhr) to my system. In fact, doesn't use of Al-Qari' al-Ali imply a "dedicated" Sakhr system? (this isn't a rhetorical question—I'd really like to know). I'd like to see a reasonably-priced (less than \$1500) Arabic OCR that runs in Arabic Windows, and has the accuracy of Al-Qari' al-Ali (or better). Apparently, this requires not only OCR but some intelligent spell-checking as well. By "intelligent" spell-checking I mean contextual analysis of syntax: there are too many Arabic words that look fine (i.e. accurately spelled) in isolation, but in context you can tell right away what's wrong. Will such a product be available soon?

Tim Buckwalter



Sakhr plans to make all their products available for MS Windows and I think Al-Qari' al-Ali was on top of their list and pretty sure it's available now.

Contact Digitek in the US. I guess you mean a grammar checker. Al-Qari' al-Ali supposedly had a grammar checker included.

Abdel-Hadi



I think Al-Qari' al-Ali is already available for MS Arabic Windows...About using another system: I use three different Windows-versions: 1) Dutch Windows 2) Arabic Windows (for workgroups, etc.) 3) Arabized English windows, Nawafidh, Sakhr for the Al-Qari' al-Ali (and no Windows 95 yet, why should I?). This all works OK, as long as every Windows is in a different subdir. The only thing is that my Dutch windows shows some garbage as group- and item-titles for some programs. (But I think this can be avoided by making sure that config.sys and autoexec.bat do not make any mention of Windows, or Windows sub-directories etc.). Of course you need some extra space on your HD, but it all works very well.

Al-Qari' al-Ali for Nawafidh certainly has no grammar checker. Just a spell-checker, as I earlier described: relation between the picture-file and the text-file. Any non-existing word (according to the program's dictionary) is shown in both. But: for example the word **mas'ul**, as written in the correct way (i. e. **hamzah** on **wāw!**), is always indicated as mis-spelled. To all our Egyptian friends: your spelling (hamzah on **yā** in **mas'ul**) is wrong.

Jan Hoogland



I have constructed a web page that includes my Ph.D. dissertation on Arabic OCR and some of my papers in postscript format. The page also includes a bibliography of Arabic OCR in three different formats. The bibliography appeared in a paper titled "Survey and bibliography of Arabic text recognition" published in Signal Processing, January 1995. You will also find a database of Arabic document images and their ground truth text, which were used in some of my experiments. You are welcome to view and download those items, but if you use them, please, cite the reference. The address of the page is <http://george.ee.washington.edu/~badr>

Badr Al-Badr

PRESENTING THE MANUSCRIPT

O. F. Akimushkin

MURAQQA'. ALBUM OF THE INDIAN AND PERSIAN MINIATURES OF THE 16—18TH CENTURIES AND THE MODELS OF THE PERSIAN CALLIGRAPHY OF THE SAME PERIOD

Muraqqa' is an interesting and rare in its repertoire collection of the Indian miniatures of the Mughal school of the 17—early 18th centuries and the Deccan school of the 17th century, as well as of the Persian miniatures of the Isfahan school of the late 16—17th centuries. It also contains the models of Persian calligraphy art.

In spite of the very high value of the wonderful miniatures of the Album — *Muraqqa'* — and the publication of 17 of them, up today its full corps was practically unknown both to the specialists of the Persian and Indian book miniature and to the students of the Persian calligraphy. The following lines represent an attempt to fill the gap.

The Album was made in the accordion (type de concertina) way: only on one side of the sheets there were miniatures or the models of the calligraphy, which were linked by pairs, four paired blocks combined (for example, ff. 14—15, 22—23, 26—27, 32—33). The size of each folio is 22.5 × 36.5 cm, 45 folios in all.

Decoration: after the mounting of the Album on every folio the preliminary marking out was made with the monochrome frame-lines, determining the bordure lines, margins and internal frames, which were to be decorated later. The artist-decorator, evidently, never started his work, as all these sheets remained unfinished except folios 1 and 45 which were decorated by the master-decorator. F. 1a: the margins of the thick glue paper against the green background with the golden ornamental painting (according to the conventional pattern) — stylized lotus and peony (*khatā'i*) flowers with the figures of the birds and phoenix in the upper part of the margins. On the frames of the sheet there is a brown bordure stripe with thin (white and green) lines, while the wide frame of the miniature is monochrome. F. 45: the margins of the thick glue paper of the rose tint with the golden ornamental painting (ac-

ording to the pattern) — spiral of the stylized flowers of peony (*khatā'i*) and lotus. Several models of the calligraphic writing (ff. 23a, 26a, 27a), the miniatures (ff. 17a, 20a, 25a, 34a), decorated with the painted frames in different water-colours and gold ornament, evidently, were once the part of other albums and afterwards, without changing, were mounted in this *Muraqqa'*.

Binding: the size is 22.5 × 36.7 cm. Because of the special way of mounting of *Muraqqa'*, the binding has only two protective covers (without the backside cover and the valve). The binding in papier mache is decorated, on the top cover, with a water-colour painting coated with lacquer. The central field is decorated with floral (roses, tulips, peonies) and plant (large-leaf branches) motifs, surrounded by a fine frame of gilded lines, against a dark red background sprinkled with gold dust. Around its edge is a large band embellished with a garland of flowers (*khatā'i*) on a black background. A fine band similar to the inside frame runs all the way around. The back cover of the binding is made and decorated in the same way. In the center of the cover, on a large leaf, is the signature of the master-decorator — "Made by the very humble Muḥammad Shafī' ibn Muḥammad Masīḥ, 1144 [1731—1732]". The two inside covers are made in the same way: a monochrome central field (gold colour), framed by a dark gold band decorated with three white lines. It is believed to come from Isfahan or, probably, Mashhad.

There is some problem about the dating of the Album. The binding dates from 1144/1731—1732. It seems to be contemporary with the Album's compilation. However, the existence of the Indian miniatures, made in the second half of the 18th century, shows that the decoration of the Album had not been completed. It was finished about the 60s of the 18th century.

The Masters of the Brush and Qalam

The Indian miniatures of two schools form a little more than a half of the *Muraqqa'* (23 folios from 45): the Mughal school of the late 16—early 18th centuries and the

Deccan school of the first half of the 17th century. The Mughal school is represented with 21 miniatures (ff. 2—5, 7, 12, 13, 19—21, 25, 18, 19, 34, 35, 37—39, 41, 44, 45),

while the Deccan school with two miniatures (ff. 30—31). No one of them has a signature, neither original, nor attributive. 9 miniatures were published (ff. 3, 4, 7, 12, 20, 28, 30, 31, 35), see T. V. Grek, *Indiiskie miniatury 16—18 vv.* (Moscow, 1971), tabs. 9, 10, 27, 7, 5, 45, 61, 62, 6.

The Persian miniatures are less numerous — there are only 14 of them (ff. 1, 6, 8—11, 16—18, 24—36, 40, 42, 43). These are the works of the Isfahan school of the late 16—late 17th centuries. 9 of them are signed by the remarkable masters. Two more miniatures are being attributed here. Artists:

1. **Šādiq beg Afshār**, who bore the literary pseudonym Šādiqī (940—1018/1533—1610) — famous artist, poet, writer and warrior. As an artist he served under Tahmasp I (r. 1524—1576). At the time of ʿAbbās I (r. 1587—1629) he was appointed the head of the Royal library, but in 1593 he was dismissed. There is one miniature undoubtedly made by him — “Portrait of Timūr Khān Turkmān” (f. 16, see *pl. 1*). “Scene at the Desolate Tavern” also, probably, belongs to him.

2. **Aqā-Ridā b. ʿAlī-Ašghar Kāshānī**, known as Ridā-yī ʿAbbāsī (d. 1044/1635) — an outstanding Persian artist, who created the whole epoch in the Iranian art, founder of the Isfahan (traditional) school of miniature of the 17th century. Two miniatures (ff. 10, 11) and also a brilliant realistic “Portrait of a Man” (f. 17, see front cover), which was made by the artist 8 months before his death, belong to him.

3. **Muhammad Yūsuf Muṣavvir** — evidently, one of the pupils of Ridā-yī ʿAbbāsī and his fellows. His miniature “The Shaykh and the Harlot” (f. 6, see *pl. 3*) was completed by him in 1068/1658.

4. **Muʿin Muṣavvir** — evidently, the most talented and prolific pupil of Ridā-yī ʿAbbāsī. His works are dated back to 1635—1707. He served under four representatives of the Safavi dynasty, including the last Sulṭān Ḥusayn (r. 1694—1722). There are two miniatures, signed by him in the *Muraqqaʿ* (ff. 9, 16), one more could be attributed to him (f. 40).

5. **ʿAlī Qulī beg Jabbādār** — the famous artist, whose works has a European touch. He worked under the reign of three last Safavides. His works are dated between A.H. 1068 and 1129 (A.D. 1657—1717). There is one miniature, belonging to him (f. 1), made in the realistic manner.

6. **Muhammad Sulṭānī** — the artist, who worked in the European trends of the Isfahan school of the 17th century. Two miniatures (ff. 42, 43) in the *Muraqqaʿ* belong to him. Without any doubt he was a contemporary of Sulṭān Ḥusayn (r. 1694—1722).

Eight Persian miniatures were published (ff. 1, 6, 10, 11, 16, 17, 42, 43), see O. F. Akimushkin, A. A. Ivanov, *Persidskie miniatury XIV—XVII vv.* (Moscow, 1968), tabs. 75, 70, 64, 63, 74, 72, 77, 78.

Calligraphy: in the *Muraqqaʿ* there are 8 models (ff. 14, 15, 22, 23, 26, 27, 32, 33) of the calligraphic writing (7 in *nastaʿliq* and one in *shikasta-yi nastaʿliq*) in Persian (7) and in Turkish (1), dated back to the middle of the 16—beginning of the 18th centuries. All of them were made in Iran. Evidently, three masters of six — ʿAbd al-Wahhāb al-Ḥusaynī al-Mashhadī (f. 32), Shāh Muḥammad Mashhadī (ff. 14, 15) and Muḥammad Ridā Tabrizī (f. 22) — were among the prominent calligraphers of their time (in the second half of the 16th century) and their names are mentioned in the narrative sources.

Structure of the Album

Miniature “Portrait of Mirzā Jalālā” (see *pl. 4*). Artist — ʿAlī Qulī (-beg Jabbādār). Isfahan school, the 70—80s of the 17th c. Annotated signature: “Him! This is the portrait of Mirzā Jalālā. The old slave, ʿAlī Qulī”. D 181, f. 1a. Size: 9.1 × 16 cm (hereafter the size of the miniature without a frame is mentioned). Water-colour, gold, silver on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, 33—34, tabl. 75; *De Bagdad à Ispahan* (Milan, 1994), 240.

Miniature “Zāhir al-Dīn Bābur on His Throne (?)”. Mughal school, 18th c. Presumably the miniature represents the founder of the Great Mughal dynasty Zāhir al-Dīn Bābur, the Timurid (d. 1530). D 181, f. 2a. Size: 5.8 × 7.7 cm. Water-colour, gold, gouache on paper.

Miniature “Portrait of Jahāngīr Keeping Flowers in His Right Hand”. Mughal school, first quarter of the 17th c. D 181, f. 3a. Size: 6.8 × 12.3 cm. Water colour, gold and gouache on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 9.

Miniature “Small Portraits of Five Different Persons and a Bud of Dog-Rose”. The miniature consists of six fragments made in different time and by different artists. Mughal school, late 17th c. It is partly executed in the technique of *siyāḥqalamī*. D 181, f. 4a. Size: 10.2 × 14.8 cm. Water-colour, gouache, Indian ink, gold on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 10.

Miniature “Portrait of Luṭf Allāh Khān”. Mughal school, early 18th c. There is an elucidatory inscription (upwards): “Luṭf Allāh Khān”. D 181, f. 5a. Size: 9.7 × 19.5 cm. Indian ink, gouache, gold on paper.

Miniature “The Shaykh and the Harlot”. Artist — Muhammad Yūsuf Muṣavvir. Isfahan school, A.H. 1068/A.D. 1658. The miniature illustrates the quatrain (*rubāʿī*) by ʿUmar Khayyām (A.D. 1040—1123), which is written in four corners of the miniature. Annotated signature (in the center up and down) on the miniature: “He! Painted by the most humble Muḥammad Yūsuf Muṣavvir in the sublime, very pure and holy [city of] Mashhad, 1068”. D 181, f. 6a. Size: 11.3 × 18.2 cm. *Siyāḥqalamī*, Indian ink, lacquer on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, 30—31, tabl. 70; *De Bagdad à Ispahan*, 242.

Miniature “Elephant Hunting”. Mughal school, late 17th c. D 181, f. 7a. Size: 11.2 × 18.8 cm. Water-colour, Indian ink, silver on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 27.

Miniature “Scene at the Desolate Tavern”. Artist — presumably, Šādiq-beg Afshār (attributed here to). Isfahan school, ca. A.D. 1600. D 181, f. 8a. Size: 13 × 19 cm. Water-colour, Indian ink, gouache and gold on paper.

Miniature “The Teacher and the Pupil”. Artist — Muʿin Muṣavvir. Isfahan school, A.H. 1085/A.D. 1674. Signature (below in the center): “Completed on Sunday 20

in the month of Rabi' II in the year 1085 for the *muraqqa'*. Mu'in Muşavvir made, practicing". D 181, f. 9a. Size: 12 × 20.2 cm. Water-colour, Indian ink, gold on paper.

Miniature "Portrait of a Man with the Chandelier". Artist — Riḍā-yi 'Abbāsi. Isfahan school, early 17th c. Signature (to the right): "Made for Muḥammad-Bāqir, the most amiable of the sons, may Allah continue his life for centuries. The humble Riḍā-yi 'Abbāsi has made". On the margins round the miniature in 14 cartouches a lyrical poem is written. Small-size *nasta'liq*. D 181, f. 10 a. Size: 8.7 × 4.8 cm. Water-colour, Indian ink on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, 28, tabl. 64.

Miniature "Darvishes' Ritual Ceremony". Artist — Riḍā-yi 'Abbāsi. Isfahan school, A.H. 1031/A.D. 1622. Signature (in the lower left corner): "This *majlis* was depicted on Friday 18 in the happy month of Shavvāl in the year 1031. The humble Riḍā-yi 'Abbāsi has executed". D 181, f. 11a. Size: 11.1 × 17.8 cm. Water-colour, Indian ink on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, 29, tabl. 63.

Miniature "Portrait of Jahāngir with the Falcon on His Left Hand". Mughal school, first third of the 17th c. D 181, f. 12a. Size: 9.6 × 16.8 cm. Water-colour, gold and Indian ink on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 7.

Miniature "Woman-Musician with the Sitar". Mughal school, late 17th c. D 181, f. 13a. Size: 4.5 × 9.9 cm. Water-colour, gold on paper.

A model of calligraphic writing (*qit'a*). Medium and small-size *nasta'liq*. Calligrapher — a famous master of the script styles *nasta'liq* and *thulth* — Shāh Muḥammad [Mashhadī], known as Vāthiqi (d. ca. 1570). The 60—70s of the 16th c. Iran (Mashhad?). 8 lines: the verses by the calligrapher. Signature: "Shāh Muḥammad executed, practicing". D 181, f. 14a. Size: 7 × 18.2 cm. Black Indian ink, tinted paper.

A model of calligraphic writing (*qit'a*). Medium-size *nasta'liq*. Calligrapher — Shāh Muḥammad [Mashhadī]. The 60—70s of the 16th c. Iran, town of Sāri. 4 lines. Signature: "This *qit'a* was written in Sāri by the humble Shāh Muḥammad". D 181, f. 15a. Size: 6.5 × 17.9 cm. Black Indian ink, tinted paper.

Miniature "Portrait of Timūr-khān Turkmān". Two artists — Šādiq-beg Afshār and Mu'in Muşavvir. Isfahan school, 17th c. The miniature was begun in A.H. 1002/A.D. 1593—1594 by Šādiq-beg Afshār (d. A.H. 1018/A.D. 1609—1610) and completed in A.H. 1095/A.D. 1683—1684 by Mu'in Muşavvir. Signature (in the lower right corner): "The portrait of the late Timūr-khān Turkmān, painted in 1002 by the late Šādiq-beg Afshār. Completed in 1095 by the miserable slave Mu'in Muşavvir. May his life be long!". D 181, f. 16a. Size: 11.6 × 19.3 cm. Water-colour. Indian ink, gold on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, 32, tabl. 74; *De Bagdad à Ispahan*, 241.

Miniature "Portrait of a Man". Artist — Riḍā-yi 'Abbāsi. Isfahan school, A.H. 1044/A.D. 1635. Annotated signature (in the top right corner): "Completed on Thursday, 6 in the month of Rabi' I in the year 1044 [August 30, 1634]. Painted by the most humble Riḍā-yi 'Abbāsi". Above and below in the cartouche there are 6 lines of the

Persian lyrical poem. Wonderful small-size *nasta'liq*. D 181, f. 17a. Size: 8.5 × 12.1 cm. Water-colour, Indian ink on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, 26—30, tabl. 62; *De Bagdad à Ispahan*, 243.

Miniature "Portrait of a Young Man". Isfahan school, second half of the 17th c. D 181, f. 18a. Size: 6.7 × 12.7 cm. Water-colour, gold on paper.

Miniature "Meeting in the Mountains". Mughal school, late 16—early 17th cc. D 181, f. 19a. Size: 6 × 11 cm. Water-colour, Indian ink on paper.

Miniature "Scene in the Harem". Mughal school, early 17th c. The miniature was once taken from some album and mounted, according to the margins with the designs of the flowers of India, in the first half of the 17th c. D 181, f. 20a. Size: 12.5 × 17.3 cm., with the margins — 19 × 26 cm. Water-colour, gold and silver on paper. Publ.: Grek, *Indiiskie miniatury*, 14, tabl. 5 (on page 52, No. 5 was erroneously attributed to the Russian National Library in St. Petersburg (call number Dorn 489).

Miniature "Jahāngir Receives Ambassadors with the Contribution". Mughal school, first half of the 17th c. Style and colours of the personages' dress and that of their head-gears allow to suppose the depicted people to be the inhabitants of the Muslim region. D 181, f. 21a. Size: 12.3 × 20.3 cm. Water-colour, silver and gold on paper.

A model of calligraphic writing (*qit'a*). Large-size *nasta'liq*. The Turkish language. Calligrapher — Muḥammad Riḍā Tabrizi (d. ca. 1595), an outstanding master of the *nasta'liq* script and *siyāhmashq*. Iran, (Tabriz?), A.H. 986/A.D. 1578—1579. Signature: "Poor Muḥammad Riḍā Tabrizi. 986". D 181, f. 22a. Size: 11 × 18.5 cm. Indian ink, paper.

A model of calligraphic writing (*qit'a*). Large-size *nasta'liq*. Calligrapher — Qavām al-Dīn Muḥammad, famous representative of the Shiraz school of calligraphy. Iran, (Shiraz?), A.H. 1077/ A.D. 1666—1667. Signature: "This was written by the poor sinner Qavām al-Dīn Muḥammad. May [Allah] forgive his sins. 1077". D 181, f. 23a. Size: 8.8 × 18.7 cm. Indian ink, golden dust on paper.

Miniature "Rustam Defeating White Dev". Isfahan school, early 17th c. Miniature borrows the subject from "Shāh-nāma" by Firdausi (d. A.D. 1010/1020). Having in mind to help Shāh Kay-Kavūs, imprisoned in Mazendaran, the hero Rustam performed seven heroic deeds. The miniature depicts the last of them. D 181, f. 24a. Size: 12.5 × 18.3 cm. Water-colour, Indian ink, gold on paper.

Miniature "An Aristocrat's Departure". Mughal school, late 16—early 17th cc. The miniature is lightly tinted with the pink and gold semis. In the center of the wide stripe, surrounding the internal frame of the miniature, there are 12 rectangular cartouches with the verses by the great Persian poet Shams al-Dīn Muḥammad Ḥāfiz (d. A.H. 791/A.D. 1389), written in wonderful small-size *nasta'liq*. The background of the frames is gilded, decorated with the gold spiral of convolvulus. D 181, f. 25a. Size: 10 × 15.5 cm. Black Indian ink, water-colour, gold on paper. The miniature is very close stylistically to the miniature of the *Muraqqa'* from The Russian National Library collection in St. Petersburg (call number Dorn 489). Cf. Grek, *Indiiskie miniatury*, 12—13, tabl. 1.

A model of calligraphic writing (*qit'a*). Large-size *nasta'liq*. Calligrapher — Turāb b. Faḍlī Anṣārī (Iṣfahānī). Signature: "The praying slave Turāb b. Faḍlī Anṣārī". The 80—90s of the 16th c. Iran—India. The text is decorated: the lines are surrounded with the points and black waved line (*tahrir-i miṣhāndān*). The gilded background has in the center herbal ornament, performed with the lighter gold and paints. At the top and at the bottom of the sheet there are two triangular cartouches of the blue background with the ornament (*islīmi-yi bargī*). D 181, f. 26a. Size: 7.3 × 17.5 cm. Water-colour, gold, Indian ink on paper.

A model of calligraphic writing (*qit'a*). Large-size *nasta'liq*. Iran, the 80—90s of the 16th c. The model is decorated masterly: the lines are framed with gold (*tarsī*), there are also the waved black stripe (*tahrir*), a flower and herbal ornament (*khatā'i* and *islīmi-yi bargī*). D 181, f. 27a. Size: 9 × 16 cm. Black Indian ink, gold and water-colour on paper.

Miniature "Sacrifice to Shiva". Mughal provincial school, first quarter of the 18th c. D 181, f. 28a. Size: 15 × 23 cm. Water-colour, gold on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 45.

Miniature "Conversation of Two Noble Indian Women in the Presence of Two Servants with Fans". Mughal school, late 17—early 18th cc. D 181, f. 29a. Size: 11.2 × 17.8 cm. Water-colour, Indian ink, gold on paper.

Miniature "Portrait of Parvīz in Full Dress". Deccan school, first quarter of the 17th c. The miniature depicts Parvīz (1589—1626), the second son of Jahāngīr (1569—1627). D 181, f. 30a. Size: 9 × 14.3 cm. Water-colour, gold on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 61.

Miniature "Portrait of Muhammad 'Adil-shāh". Deccan school, middle of the 17th c. The miniature depicts Muḥammad 'Adil-shāh, ruler of Bijapur (r. 1626—1658). D 181, f. 31a. Size: 3.7 × 5 cm. Water-colour, gold on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 62.

A model of calligraphic writing (*qit'a*). Medium-size *nasta'liq*. Calligrapher — 'Abd al-Wahhāb al-Husaynī al-Mashhadī (d. ca. 1575). Iran (Mashhad), the 40—60s of the 16th c. This *qit'a* contains four lines from the poem by Badr al-Dīn Hilālī (killed in A.H. 936/A.D. 1529) "Ṣifāt al-'Ashiqīn". Signature: "Abd al-Wahhāb al-Husaynī al-Mashhadī, may Allah forgive his sins and guilt". D 181, f. 32a. Size: 9.6 × 18.7 cm. Black Indian ink, paper.

A model of calligraphic writing (*qit'a*). *Shikasta-yi nasta'liq*. Calligrapher — Muḥammad Mu'min al-Qumī. Isfahan, 14 Dhū-l-qa'da 1116/March 10, 1705. Seven lines and monogram: "It was made in the paradise like Isfahan, the chandelier of the state, 14 in the holy month of Dhū-l-qa'da. Muḥammad Mu'min has written this. 1116". D 181, f. 33a. Size: 10.2 × 18.2 cm. Black Indian ink, grey tinted paper with gold semis.

Miniature "Portrait of a Noble Man". Mughal school, last quarter of the 17th c. D 181, f. 34a. Size: 8.2 × 14.8 cm. Water-colour, Indian ink and gold on paper.

Miniature "Portrait of Akbar". Mughal School, middle of the 18th c. The miniature depicts Great Mughal Emperor Jalāl al-Dīn Akbar I (r. 1556—1605). D 181, f. 35a. Size: 9 × 14.6 cm. Water-colour, gold on paper. Publ.: Grek, *Indiiskie miniatury*, tabl. 6.

Miniature "The Grades of Getting Drunk with the Love to the Absolute Truth". This wonderfully made miniature, probably, suffered the influence of the works by Persian artist Muḥammadi Muṣavvir, who worked in Herat

and Mashhad in the 70—90s of the 16th c. Evidently, Isfahan school, the 10—20s of the 17th c. D 181, f. 36a. Size: 14.5 × 25.5 cm. Water-colour, Indian ink, gold on paper. See A. K. Coomaraswamy, *Les miniatures orientales de la collection Goloubew au Museum of Fine Arts de Boston* (Paris—Bruxelles, 1929), 35, No. 48, pl. XXVI; B. Gray, *La Peinture Persane* (Genève, 1971), 157 (in colour); B. W. Robinson, "Muḥammadi and the Khurasan style", *Iran*, vol. XXX (1992), 17—29.

Miniature "Dinner in the Royal Tent at Hunting". Mughal school, first quarter of the 17th c. Probably, this multi-figured composition had not been finished, only gold was put on. D 181, f. 37a. Size: 13.5 × 23.2 cm. Gouache, Indian ink and gold on paper.

Miniature "Portrait of Shāh 'Ālam I Bahādur-shāh". Mughal school, first quarter of the 18th c. The miniature is incomplete. At the top of it there is a signature: "Shāh 'Ālam Bahādur-shāh, the governor and the warrior for faith", i. e. the Great Mughal Emperor Shāh 'Ālam (r. 1707—1712). D 181, f. 38a. Size: 13 × 21.2 cm. Water-colour, Indian ink on paper.

Miniature "Portrait of Sulṭān Maḥmūd". Mughal school, late 17—early 18th c. The miniature was enlarged at the bottom, but not completed. At the top there is a signature: "Portrait of Sulṭān Maḥmūd, the son of His Majesty 'Ālamgīr", i. e. son of the Great Mughal Emperor Muḥyī al-Dīn Awrangzīb (r. 1658—1707), who had a honourable name (*laqab*) 'Ālamgīr (the Conqueror of the World). D 181, f. 39a. Size: 13 × 20 cm. Water-colour, Indian ink, gold on paper.

Miniature "Intimate Meal in the Company of Musicians". Artist — presumably, Mu'in Muṣavvir (attributed here to). Isfahan school, early 17th c. D 181, f. 40a. Size: 12.5 × 17.5 cm. Water-colour, Indian ink and gold on paper.

Miniature "Portrait of a Noble Man with a Page in White Dress". Mughal school, late 17th c. During the mounting of the Album the miniature was enlarged (at its top and bottom). D 181, f. 41a. Size: 12 × 22.8 cm. Water-colour, silver, gold on paper.

Miniature "Portrait of a Noble Man with a Stick in His Left Hand". Artist — Muḥammad Sulṭānī. Isfahan school, late 17th c. Signature (above): "Muḥammad Sulṭānī has painted [that]". D 181, f. 42a. Size: 19 × 28.5 cm. Water-colour, silver and gold on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, tabl. 77; *De Bagdad à Ispahan*, 245.

Miniature "Portrait of a Noble Man". Artist — Muḥammad [Sulṭānī]. Isfahan school, A.H. 1109/A.D. 1697—1698. Signature: "Completed for the honourable vice-regent [by] the most insignificant of the slaves Muḥammad. 1109". D 181, f. 43a. Size: 16.2 × 10.8 cm. Water-colour, silver, gold on paper. Publ.: Akimushkin, Ivanov, *Persidskie miniatury*, tabl. 78.

Miniature "Old Awrangzīb on the Throne". Mughal school, early 18th c. D 181, f. 44a. Size: 19.5 × 26.5 cm. Water-colour, silver and gold on paper.

Miniature "Portrait of a Noble Courtier". Mughal school, middle of the 17th c. D 181, f. 45a. Size: 11 × 19 cm. Water colour, silver and gold on paper.

The *Muraqqa'* (Album) preserving in St. Petersburg was once an item in the collection of A. Ya. Italinsky (1743—1827), Russian physician and diplomat, ambassa-

dor to the Ottoman Turkey. In 1829 *Muraqqa'* was brought with other 90 Muslim manuscripts of his collection to the Educational Department of the Asiatic branch of the Ministry of Foreign Affairs of the Russian Empire, where it was kept till 1919. After that the whole collection of manuscripts of the Educational Department was transferred to the Asiatic Museum of the Russian Academy of Sciences (now St. Petersburg Branch of the Institute of the Oriental Studies of the Russian Academy of Sciences), where one can find it today (call number D 181).

Bibliography: J. Hammer, [“Notice”], *Bibliotheca Italiana* (Milan, 1828), vol. XLIX, p. 22; V. R. Rosen, *Les manuscrits persans de l'Institut des langues orientales du Ministère des Affaires Etrangères* (Saint-Petersbourg,

1886), No. 131, p. 322; *Persidskie i tadjikske rukopisi Instituta Narodov Azii AN SSSR (Kratkii alfavitnyi katalog)* (Persian and Tadjik Manuscripts of the Institute of Oriental Studies, USSR Academy of Sciences. The Concise Alphabetic Catalogue), i (Moscow, 1964), No. 4034. O. F. Akimushkin, A. A. Ivanov, *Persidskie miniatiury XIV—XVII vv. v leningradskikh sobraniakh* (Persian Miniatures of the 14—17th Centuries in the Collections of Leningrad) (Moscow, 1968); *De Bagdad à Ispahan. Manuscrits islamiques de la Filiale de Saint-Petersbourg de l'Institut d'Etudes orientales, Académie des Sciences de Russie. Musée du Petit Palais. 14 octobre 1994—8 janvier 1995* (Milan, 1994).

Illustrations

Front cover:

Fol. 17a. Portrait of a Man by Riḍā-yi ‘Abbāsi, 11.8 × 8.2 cm.

Back cover:

Plate 1. Fol. 16a. Portrait of Timūr Khān Turkmān by Ṣādiqī beg Afshār, 19.3 × 11.6 cm.

Plate 2. Fol. 36a. The Darvishes Picnic in the Mountains. Probably Isfahan school, 25.5 × 14.5 cm.

Plate 3. Fol. 6a. The Shaykh and the Harlot by Muḥammad Yūsuf Musavvir, 18.2 × 11.3 cm.

Plate 4. Fol. 1a. Portrait of Mirzā Jalālā by ‘Alī Qulī beg Jabbādār, 16.0 × 9.1 cm.

BOOK AND SOFTWARE REVIEW

'Ajā'ib al-Dunyā (Chudesa mira). Kriticheskii tekst, perevod s persidskogo, vvedenie, komentarii i ukazateli L. P. Smirnovoi. Moskva: Nauka, 1993, 540 str. — Pamiatniki pis'mennosti Vostoka, LXXXIII.

'Ajā'ib al-Dunyā (Wonders of the World). Critical text, Russian translation from Persian, introduction, commentary and indices by L. P. Smirnova. Moscow: Nauka Publishing House, 1993, 540 pp. — Literary Monuments of the Orient, LXXXIII.

The publication is a volume of the well-known series. The work under review comprises preface, critical text, Russian translation of the text, commentary to the translation and supplements including: 1. Abbreviations; 2. Bibliography; 3. Index of geographical and ethnic names; 4. Index of persons; 5. Index of subjects; 6. Abbreviations used in the Persian text; 7. Summary.

In the detailed preface L. Smirnova considers the history of the genre of *'ajā'ib* and of the monument itself, gives a study of the paleography and philological peculiarities of its manuscripts. She also considers the problems of the authorship and of the dating, as well as of the sources of the monument. Basing on the textological comparison of three manuscripts of *'ajā'ib* (the manuscript preserving in the Manuscript Collection of the St. Petersburg branch of the Institute of Oriental Studies of the

Russian Academy of Sciences, call No. A 253, ff. 87a—228b; the manuscript from the collection of B. G. Browne at the Cambridge University Oriental Library, call No. G 11 (12), ff. 44b—72b and the copy from the depositary of the Library of the Majlis in Tehran (37503), which came from the private collection of Mālik al-Shu'arā Dr. M. T. Bahar), L. Smirnova comes to the conclusion that the Cambridge and Tehran copies are more close to the author's text than the manuscripts from St. Petersburg.

The monument belongs to the genre which was very popular in the Medieval Muslim World. It includes the real scientific information on geography, history, cosmography, mineralogy, ethnology, though mixed with legends and myths. The tradition of creating of such books goes back probably to the Zoroastrian Pahlavi literature like *Shahristaniha i Eran* (Towns of Iran), *Abdiha ud Sahigiha i Sagestan* (Miracles and Wonders of Sistan) and others.

The publication of the critical text of *'Ajā'ib al-Dunyā* and the translation, provided with the extensive bibliography, commentaries of utmost interest and indices, is an important and valuable contribution to the study of Persian culture and language, as well as of the Near and Middle Eastern history and culture in general. It is also helpful for the study of historical and cultural traditions of Georgia, Armenia, Azerbaijan and Turkestan.

F. Abdullaeva

Avesta. Izbrannye gimny iz Videvdata. Perevod s avestii-skogo Ivana Steblin-Kamenskogo. Moskva, 1993, 207 str.

Avesta. Selected Hymns from Videvdad. Russian translation from Avestan by I. Steblin-Kamensky. Moscow, 1993, 207 pp.

It was almost two years ago when the fullest Russian translation of a part of Avestan text was published, which contains the hymns to eight gods and goddesses of Old Iranian pantheon (Ahura Mazda, Ardivisura, Khurshed, Tishtriya, Mitra, Veretragna, Ashi, Khwarno) and a prose passage from Videvdad (fragard 2). Thus the lacuna in the study and publishing of the Holy Books of the world-wide religions in Russia is filling up.

The publication consists of the rather short preface, the translation with commentaries and a general index of personal names, geographical names and terms, met in the text and in the commentaries.

In spite of the modest size of the preface by Prof. Steblin-Kamensky one can find here a funded survey, dedicated not only to the text of Avesta itself — the Holy Book of the Zoroastrians, but also a wealth of information about the Prophet Zaratushra (who lived at the end of the second half and at the beginning of the first millennium B.C.), the history and the religion in general.

The most difficult task of the work was the adequate translation of the sayings by Zaratushra which can be translated and interpreted in quite different ways. Old Iranians considered the art of writing as created by the Devil — Angra Mainyu. That is why the fixation of the Holy texts in a written form was not allowed. All of them were to be known by heart and transmitted from generation to generation of priests orally. From the days of the emergence of Zoroastrianism, when it was the religion of cattle-breeders with an elementary organized cult structure during almost three millenniums of its history, Zoroastrianism turned into a very elaborate official religion of great Empires, such as the Empires of Sassanids, of Parthian Ar-

shakids with its numerous priests, a great number of temples and specific way of liturgy, as well as with many every day life restrictions. It influenced greatly other religions. One can find the traces of it in Judaism, Christianity, Islam and Buddhism. At the same time Zoroastrism suffered the influence of other religions. One of the result of it — writing down the Avestan text in the Pahlavi script — should be taken into account.

The translation of such kind of the text could be done by a specialist not only in the field of the history of Zoroastrian religion, but by a historian, ethnologist and philologist with the really encyclopedic knowledge of Indo-European dead and living languages. That is I. M. Steblin-Kamensky who is, fortunately, such kind of a specialist. His Iranian studies are widely known all over the world and we could not miss a chance to congratulate Prof. I. M. Steblin-Kamensky, who has recently been appointed the Dean of the Oriental Faculty of the St. Petersburg University (after Academician M. V. Bogolyubov having retired), on the occasion of his 50th anniversary.

The earliest manuscript of Avesta dates back to 1288. The text of Avesta is still considered to be one of the most difficult ever created in Indo-European languages.

The translation by I. M. Steblin-Kamensky follows not only the sense and exact meaning of the original text, it also reproduces the very structure of the Hymns. Most of the Yashts (Hymns) were created in the form of 7—9 syllable verses (sometimes 4—13 syllable), which was carefully preserved by Prof. Steblin-Kamensky. The poetic style of the translation is very clear and natural. It is absolutely relevant to the ancient Holy Hymns.

The work under review represents an important contribution to the Avestan and Iranian studies in general. Both the common readers and the students of the East will certainly enjoy Prof. Steblin-Kamensky's compact and integrating presentation of the ancient text. The next part of the Avestan translation by Steblin-Kamensky — the Gathas (the Songs of Zaratusht) — is expected in the near future.

F. Abdullaeva

ADOBE ACROBAT 2.0 AS A MEDIUM FOR ELECTRONIC PUBLISHING IN ARABIC STUDIES (WINDOWS VERSION)

Producer: *Adobe Systems Incorporated*

System requirements:

- 386- or 486-based personal computer (or Pentium)
- Microsoft Windows 3.1 or greater
- 4 Mb of RAM for Acrobat Exchange and Acrobat Reader
- 8 Mb of RAM for Acrobat Distiller

Method of testing:

- computer used
Compaq DeskPro XL 590, CPU Pentium 90,
16 Mb RAM
- platform
Microsoft Windows 3.1
Microsoft Windows 3.1 with Arabic support
("bi-l-i'timad")
- texts and PostScript files produced with
Microsoft Word for Windows 6.0a
Microsoft Arabic Word for Windows 6.0a
NISUS 3.4 (Macintosh)

The Adobe Acrobat Pro 2.0 package comes with Adobe Reader 2.0, Adobe Exchange 2.0 (PDFWriter), and Adobe Distiller 2.0, together with a limited purpose version of Adobe Type Manager 3.01. An upgrade of the Reader (version 2.1) is available from Adobe Systems (<ftp://ftp.adobe.com> or <http://www.adobe.com>).

The Windows version of the Adobe Acrobat family of programs performs essentially the same as the Macintosh version described in an earlier review (*Manuscripta Orientalia* 1/2, <http://www.hf-fak.uib.no/i/midtspraak/acrobat.htm>). In general PDF (portable document format) files retain all the formatting and detail of hard copy and can handle both Latin diacritics and Arabic script, although there are occasional problems with screen fonts, and image quality must be reduced if files are to be kept to a manageable size. PDF files can be produced directly from word

processing programs by PDFWriter, which is part of Acrobat Exchange. They may also be generated from PostScript files by Acrobat Distiller, including files containing Arabic fonts. PostScript files can be created from most word processor or page layout programs or by using the printer option offered by Acrobat Distiller Assistant. Acrobat Reader, which allows viewing on screen, simple searching, and printing out, is available free. Acrobat Exchange permits the inclusion in a PDF file of hypertext links to other files, including sound and video clips. Exchange also provides sophisticated searching techniques (see below on Adobe Catalog). Compared to word processor files, PDF files tend to be rather large, but they can be considerably reduced, often to little more than the size of the word processor original, by the exclusion of Type 1 (PostScript) fonts.

Production of PDF Files. It is possible to render Arabic TrueType fonts with PDFWriter on the PC platform using US Word in US Windows and to minimize file size by excluding the fonts from the file, but the Arabic text will not be properly coded on the Macintosh platform. In Arabic Windows, moreover, PDFWriter does not render Arabic script produced by Arabic Word for Windows. Since PDFWriter is therefore generally not appropriate for our purposes, the basic task of interest here is the production of PDF files from PostScript files with Acrobat Distiller. Once a PDF file is produced, Adobe Reader or Exchange, with minor exceptions, can handle it properly.

The way Acrobat Distiller works with TrueType fonts is unfortunately far from satisfactory. TrueType fonts are converted to Type 3 (bitmap) fonts, and cannot be excluded from PDF files as is possible with Type 1 fonts, or with TrueType fonts when using PDFWriter. Files containing Type 3 fonts can be five times, or more, the size of equivalent files based on Type 1 fonts. Moreover, the quality of the screen display of TrueType fonts is inferior to that of Type 1 fonts. Text is sometimes almost illegible, a good result often being obtained only when zooming up to 400% or more. The printout quality is very good, however, being in every respect equal to a printout from the program in which the document was originally produced. The problems with TrueType fonts are particularly significant for the publication of documents containing Arabic script, since all the Arabic fonts in Microsoft Arabic Windows, at least in the version we are using, are TrueType fonts. With Arabic TrueType files from Arabic Windows Common Ground's Digital Paper format gives considerably better results than Acrobat Distiller. The files are smaller, and the screen display is much easier to read.

Another problem seems to lie in the specification of the Arabic TrueType fonts in PostScript files from Microsoft Arabic Windows. The TrueType fonts in the PostScript files are specified only at the end of the file, and interpreting them on occasion seems to cause difficulties for Acrobat Distiller. Quite often they are interpreted as something else (usually Type 1 Courier), or as an offending command. In the latter case no PDF file is created. This problem may at times be connected with combining different sizes of letters.

Sometimes in the case of a longer document (16 pages, 16 point type) only the odd pages were produced in Arabic script, while all the even pages were rendered in Courier. A similar thing occasionally happened when combining Arabic and Latin script, most of the Arabic parts being reproduced normally, while some were rendered as Courier. We found no regular pattern here.

Acrobat Distiller for Windows can be used to make PDF files from PostScript files produced on a Macintosh, and the size of the files can be reduced by eliminating Macintosh Type 1 fonts, including diacritic fonts. It is sufficient to write the names of the fonts in the "Never Embed" list and disable the "Embed All Fonts" option. Unfortunately Arabic fonts cannot be excluded. The results are very good on the screen and excellent when printed out. The Windows version of Common Ground (2.01), which is the only version currently available, is noticeably inferior to Distiller in rendering PostScript files, whatever the original platform.

Displaying of PDF Files. As on the Macintosh, Type 1 fonts must be uninstalled when browsing files that contain them, whether or not they are embedded. The Reader seems to look for the fonts installed in the system and use them for browsing, which can lead to conflicts between them and the fonts in the PDF file. With the exception of the poor screen display of TrueType fonts mentioned above, there are no great problems with browsing in US Windows. In Arabic Windows, however, one must remember to disable Adobe Type Manager, if it is installed, otherwise Acrobat Reader may fail to function, as it did on the machine we were using.

Searching. Acrobat Reader offers non-problematic simple searching ("Find") in lower ASCII Latin text. In Arabic, however, while there are some indications that searching may be possible in a future version of the Reader, this is not the case at the present time. Certain differences between apparently identical words result in the next occurrence of the word not always being found. As in the Macintosh version, there are problems searching for upper ASCII Latin diacritics, and the various Arabic glyphs are not reduced to their least common denominators, which renders meaningful searching almost impossible. Because of the new Acrobat encoding, the keyboard cannot be relied upon to type upper ASCII characters into the Find window. The case is even more complicated with Arabic glyphs. Copying and pasting produces quicker results than the keyboard, but is still not entirely reliable. The simple Find function in Acrobat Exchange performs similarly.

Copying and pasting text. Lower ASCII text may be copied from PDF files opened in Reader or Exchange and pasted into word processor files without difficulty, except that line breaks become paragraph breaks. Copying and pasting of upper ASCII Latin and of Arabic characters gives chaotic results, however. Arabic text originally from Arabic Word for Windows, for example, cannot be pasted back into that program with success.

Adobe catalog. This tool, which includes software licensed from Verity, Inc., and which is sold separately from the Adobe Acrobat Pro package that contains Distiller, Exchange and Reader, creates indexes of the PDF files stored in designated directories. The updating of indexes can be automated and set to run at specific times, every hour, for example. Once the indexes are created, it is possible in Exchange to form various types of queries, including Boolean, stem, and semantically related word searches. Indexing is a very simple and relatively fast process. With a large amount of data, however, updating the index at close intervals could make it difficult to work simultaneously with other programs. Searching for upper ASCII Latin and Arabic, unfortunately, has all the shortcomings mentioned in connection with the Find function in Reader and Exchange. With Arabic, moreover, one must search for glyphs rather than characters, and even words that seem to be identical will not necessarily be found in a search.

Conclusion. Despite the problems mentioned above, Adobe Acrobat remains a superior tool for electronic publishing, whether online or offline. The primary advantage in the Acrobat system seems to be the use of real fonts

which permit real text searches but which do not have to be embedded in all files in a given series. For browsing it is sufficient to open first a single file which has the fonts used embedded. Subsequent files opened do not have to have the fonts embedded and can thus be many times smaller. It is reasonable to assume that the present problems with the encoding of upper ASCII and Arabic glyphs

will be solved with the introduction of a more universal coding system.

This review is preserved in electronic form and in hard copy in the Archive of Electronic Publications of the Section for Middle Eastern Languages and Cultures, University of Bergen (<http://www.hf-fak.uib.no/i/midtspraak/aep.htm>).

J. Bell

Persian and Tajik Manuscripts

of the Russian Academy of Sciences
Institute of Oriental Studies: Short Catalog

Reprint with Addenda and Translations in 2 volumes

Edited by **N. D. Miklukho-Maklai**, with Corrigenda and Addenda by **O. F. Akimushkin**

The first edition of the Short Catalog of **Persian and Tajik Manuscripts** of the Institute of Oriental Studies in St. Petersburg, Russia was originally published in 1964 in Russian and Persian and is now out of print. This reprint edition, consisting of an alphabetical list of 4912 different titles found in some 3100 manuscripts, has been enhanced by the addition of the following new information:

- **ADDENDA SECTION** consisting of 122 new entries describing the manuscripts new to the collection since 1964
- **ENGLISH TRANSLATIONS** of the introduction, table of contents, and subject index
- **CORRIGENDA SECTION** correcting errors from the 1964 edition

Persian and Tajik Manuscripts of the Russian Academy of Sciences, Institute of Oriental Studies: Short Catalog, 2 volumes. Reprint edition in library binding, printed on acid-free paper.

Prepublication price through November 30, 1995, both volumes \$ 180

List price, both volumes \$ 220

NORMAN ROSS PUBLISHING INC.

550 West 58th Street, New York, NY 10019 USA

212/765-8200 • 800/648-8850 • FAX: 212/765-2595 • E-MAIL: nross@igc.apc.org

PO BOX 111, St. Petersburg, Russia, 198152
E-mail: info@petvost.spb.su

ST. PETERSBURG CENTRE
FOR ORJENTAL STUDJES PERJODJICALS



St. Petersburg Journal of Oriental Studies (ISSN 0869-8392)

Chief editor I. Alimov. Editors S. Suhachev, O. Trofimova. Co-editors S. West, R. Ames, M. Rodionov.

- **Vol. 1** St. Petersburg, 1992, 398 pp. ● **Vol. 2** St. Petersburg, 1992, 398 pp.
- **Vol. 3** St. Petersburg, 1993, 456 pp. ● **Vol. 4** St. Petersburg, 1993, 492 pp.
- **Vol. 5** St. Petersburg, 1994, 632 pp. ● **Vol. 6** St. Petersburg, 1994, 704 pp.
- **Vol. 7** St. Petersburg, 1995, 656 pp.

The edition gives a broad panorama of a present-day of ancient and mediaeval East research. This periodical publishing articles, translations, bibliographies, reviews, archive materials written by various specialists has no analogues in Russia.
It is issued at least two times a year.

Kunstammer. Ethnographic papers (ISSN 0869-8384)

Co-editors M. Albedyll, M. Rodionov, I. Alimov, S. Suhachev

- **Vol. 1** St. Petersburg, 1993, 232 pp. ● **Vol. 2–3** St. Petersburg, 1993, 492 pp.
- **Vol. 4** St. Petersburg, 1994, 392 pp. ● **Vol. 5–6** St. Petersburg, 1994, 496 pp.
- **Vol. 7** St. Petersburg, 1995, 272 pp. ● **Vol. 8–9** St. Petersburg, 1994, 512 pp.

This journal set up in 1993 is a new periodical devoted to traditional culture of the peoples of the world. The journal publishes materials about Kunstammer, the oldest Russian museum, and other museum stocks of St. Petersburg, as well as research articles dedicated to every-day life and traditions of the peoples of the world.

Please send your book-orders to:

OTTO HARRASSOWITZ. BUCHHANDLUNG
Tannusstraße 5, Postfach 29 29, Wiesbaden, D-65019.

KUBON UND SAGNER. BUCHEXPOR-IMPORT GmbH.
Heßstraße 39/41, Postfach 34 01 08, München, D-80328.

AUTHORS

Dr. *Firuzia I. Abdullaeva* — Professor of Persian at the Oriental faculty of the St. Petersburg State University, specialist in the Iranian Philology and Islamic culture. Author of a series of works in the field.

Prof. Dr. *Oleg F. Akimushkin* — Head of the Dpt. of the Middle East and Group of the Kurdish studies of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, specialist in the history of Persian culture, history of Iran, Central Asia and Eastern Turkestan, author of a series of monographs and articles.

Prof. *Joseph N. Bell* — Professor of Arabic and Head of Section for Middle Eastern Languages and Cultures, University of Bergen, Bergen, Norway. Specialist in Muslim intellectual history and Sufism. Author of several books and a series of articles in the field.

Dr. *Ksenia B. Kepping* — Senior Researcher of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, specialist in Tangut language and culture, author of a series of monographs and articles.

Mr. *Alexander S. Matveev* — post graduate student of the Oriental Faculty of St. Petersburg State University and special student of Prof. Joseph Bell from University of Bergen. Specializes in Arabic medieval history, author of several articles in the field.

Prof. Dr. *Lev N. Menshikov* — Head of the Group of Far Eastern Textology of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, specialist in the history of Chinese literature and culture, founder of Russian school of Dunhuang studies and author of a series of monographs and articles.

Dr. *Tatiana Pang* — Researcher of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, author of a series of works devoted to the Manchu manuscript heritage, culture and philology.

Dr. *Efim A. Rezvan* — Research Director of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, specialist in Arabic and Islamic studies, author of the monographs and series of articles devoted to the history of the Russian-arabic relations, history of Islam and computer approaches in the field.

Dr. *Paul Roochnik* — Arabic Computational Linguist at Applications Technology (AppTek) Inc., McLean, Virginia, USA. Founder and moderator of ITISALAT, the Internet Forum on Arabic Computing. Co-founder and secretary of The Association for Arabic Computing in North America. Specialist in Arabic Machine Translation. Author of a series of articles on the computerized parsing of Arabic sentences as well as a series of articles on the computerized treatment of Arabic personal names.

Dr. *Tursun I. Sultanov* — Senior Researcher of the St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, specialist in medieval handwritten tradition and culture of Turkish peoples, author of the monographs and series of articles in the field.

Dr. *Petr Zemanek* — Deputy Head, Institute of Ancient Near Eastern Studies, Charles University, Prague, Czech Republic. Specializes in Arabic and Semitic linguistics, especially in phonology, and computer-aided analysis of Arabic and other Semitic languages.

Notes to Contributors

Manuscripts must be written in English.

Manuscripts must be clearly typewritten with numbered pages, double linespacing and wide margins throughout. Italic and bold typeface should be avoided. Use underlining where text is to be italicized. The title should be as brief and informative as possible. The institute at which the work has been done should be indicated at the head of each paper. Authors are requested to include their e-mail address if one is available.

Submissions

Manuscripts should be sent in duplicate to the Editor-in-Chief: Professor Dr. Yuri A. Petrosyan, St. Petersburg Branch of the Institute of Oriental Studies, Russian Academy of Sciences, 18 Dvortzovaya nab., 191186, Saint-Petersburg, Russia, E-mail: orient@ieos.spb.su



Plate 1



Plate 2

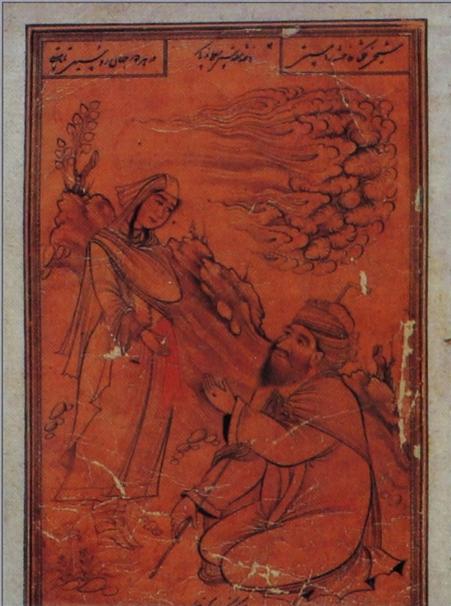


Plate 3



Plate 4