# CONTENTS

Front cover:

Fragment of the drawing demonstrating acupuncture points which should be applied when healing back pain
and a half-body paralysis. *Sabsire sûiha sindara ferguwecuke argan*, manuscript B 92 mss in the collection
of the St. Petersburg Branch of the Institute of Oriental Studies, illustration 13, fol. 42, 27.0×46.3 cm.

Back cover:

Drawing of a male figure with acupuncture points which should be applied when healing child's night crying
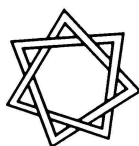and tooth-ache. The same manuscript, illustration 1, fol. 28, 27.0×46.3 cm.
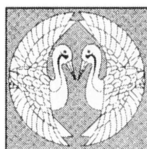
Manuscripta Orientalia

International Journal for Oriental Manuscript Research

ORIENTAL MANUSCRIPTS
AND NEW INFORMATION TECHNOLOGIES

E. A. Rezvan, N. S. Kondybaev.

THE ENTRAP SOFTWARE: TEST RESULTS

In the article published in *Manuscripta Orientalia* in September 1996 [1], we declared our intention to publish test results for the ENTRAP software, a program intended for the semi-automatic analysis of hand-written script. A large Qur'ānic fragment (call number E20) [2], held at the St. Petersburg Branch of the Institute of Oriental Studies, provided us with a good material to test the software suggested. This fragment, a vertical-format codex (34.0×52.5 cm), is undoubtedly one of the most valuable early copies of the Qur'ān to have reached us. It consists of 81 folios; the text is written on a high-quality parchment and contains 39.3% of the whole text of the Qur'ān. In all, the surviving folios comprise (in full or partly) the text of forty-four *sūra*s (2—11, 20, 24—39, 43—58, 70 and 71). The text of twenty-two of these is complete (*sūra*s 10, 25, 31—34, 36—38, 45—57).

The surviving fragment reveals the hands of two copyists (we designate their hands as A and B), who divided their work into two equal parts. The first transcribed *sūra*s 2—11, which makes up the first half of the Qur'ān's text, and the other — the text beginning with *sūra* 20 and further. It is, however, possible that the second scribe began his work with *sūra* 17, the beginning of which corresponds to the beginning of *juz'* 15, which starts approximately in the middle of the Qur'ānic text. The number of lines per page varies widely (hand A — from 23 to 31 lines; hand B — from 21 to 26), as does the size of letters. No sign of preliminary ruling can be seen.

Both hands of the fragment have common features with the style of the handwriting known from written texts from Cairo, Damascus, Ṣan'ā', as well as from the al-Ṭā'if's inscription dated to A.D. 677—78. The usual designation of this style is "late Ḥijāzī". Taking into account Fr. Déroche's classification, both hands can be considered as corresponding on the whole to styles B I and B II of the "early 'Abbāsid scripts" [3]. The fashion of writing medial *jīm/ḥā'/khā'*, however, corresponds mainly to style A I, according to the same classification. On the whole, the writing in our manuscript reflects a transitive phase of Arabic script development, from Ḥijāzī to later writing styles. *Alif* is usually written perpendicular to the line, although it is sometimes slightly inclined to the right. The same is with the vertical stroke of *ṭā'* and *lām*. Hand B, surer and more professional,

is characterised by a distinctly rounded end of final *jīm/ḥa'/khā'*, *'ayn* and *ghayn*, which distinguishes hand B from hand A.

The manuscript of the Qur'ān dates to the late eighth—early ninth century. It seems to demonstrate the mature stage of the development of one of the two early written styles of copying Qur'āns, closely linked to Northwest Arabia and to the region of Syrian border. A large number of the manuscript's orthographic and palaeographic features links it with the Ḥijāzī manuscripts discovered in Ṣan'ā'. They belong to the Qur'ān type designated by Estelle Whellan as "type 2" [4].

The establishing of even the smallest variance in hands can be of use for attributing Muslim manuscripts. This consideration explains our interest in the ENTRAP software. The question was whether the ENTRAP software can be used to distinguish hands' variance properly. To answer this question, hands A and B of our manuscript, very close to each other, were employed.

For our analysis, we chose different positions (from five to nine) of Arabic letters — initial, medial and final for *'ayn/ghayn* and *alif*, final — for *nūn* and *ṭā'*, medial — for *mīm*, initial — for *jīm/ḥā'/khā'*, medial — for *ḥā'*, final — for *qāf*, and the ligature *alif/lām* as they are written by the scribes of the Qur'ān. The letters were chosen at random. The analysis was conducted as follows: (i) letter images were borrowed from the scanned variant of the manuscript's text; (ii) a group of parameters was obtained for every symbol for creating a statistical model; (iii) variances for two groups of symbols, representing hand A and hand B, were established and analysed; (iv) hypothesis based on the statistical model employed was tested.

The results of our calculations are shown in *Table 1*. The analysis was conducted with the aid of statistical methods (dispersion analysis was used). We tested hypothesis H_0 that the mean values (expected values) of the measured parameters for a certain symbol will coincide with the values obtained for the same parameters of the same symbol in each of the two groups (hand A and hand B).

This means that if hypothesis H_0 fails for even one of the parameters analysed, the expected values are not equal. In other words, the probability distribution of values of a parameter is divided by the sum of the two probability dis-

tributions. These groups are different for the writing of the same symbol because they belong to different hands.

The statistical criterion was formulated in the following manner: if $F_{r-1,n-r} > C$, then hypothesis H_0 on the coincidence of the mean parameter values from different groups fails, where

F — distribution with r-1, n-r degrees of freedom

r — number of groups

n — number of copies of any symbol

C — constant from table of F-distribution under level of test value.

Mean values for all parameters of the symbol for each group are listed in *Table 2*. Three parameters were analysed:

As — degree of possible distortion

En — entropy

Ma — expected result

The test of statistical hypothesis H_0 was conducted for the following level of criterion significance: $\alpha = 0.05$, $C = 5.32$ and for $\alpha = 0.01$, $C = 11.26$ and $r = 2.0$. Results are shown in *Table 3*. Five symbols met our statistical criterion. Thanks to these symbols, we can distinguish hand A from hand B with a reasonable level of confidence.

The approach described in our previous article and realised by means of the ENTRAP software reveals the difference between very similar hands. It gives us the opportunity to identify hand-written symbols.

As is seen from *Table 2*, the approach described above can be of use for solving the problem of symbol classification as part of the task of automatic optical character recognition for hand-written texts. A statistical model employed enables us to see consistent separation of symbols by classes. Reflected in *Table 2* a powerful ICR (Intelligent Character Recognition) system could be based upon this approach.

Table 1

The results of computation

1.	Initial 'ayn				2.	Alif			
	A		B			A		B	
	symbol	parameter	symbol	parameter	symbol		parameter	symbol	parameter
As		186.75		185.806			189.252		189.05
En		7.319		7.664			7.078		7.006
Ma		29.825		32.574			33.528		33.537
As		186.356		184.455			189.503		189.307
En		6.989		7.33			6.976		7.03
Ma		28.575		30.631			33.065		33.636
As		185.513		186.583			188.593		189.606
En		7.416		7.346			7.442		6.876
Ma		30.508		28.939			32.858		32.438
As		184.957		186.21			189.708		189.279
En		7.511		7.514			7.06		7.076
Ma		30.883		31.74			32.652		33.889
As		186.194		186.421			189.458		189.271
En		7.503		7.309			7.055		7.052
Ma		30.234		30.808			34.073		33.752

3.	Final *nūn*				4.	Middle *'ayn*			
	A		B			A		B	
	symbol	parameter	symbol	parameter		symbol	parameter	symbol	parameter
As En Ma		188.419 7.48 35.371		188.94 6.909 31.635	As En Ma		180.264 8.126 37.332		177.062 7.938 41.906
As En Ma		187.611 7.634 35.195		187.891 7.285 32.61	As En Ma		185.442 8.06 32.964		180.92 8.107 40.422
As En Ma		188.717 7.526 34.356		189.294 7.504 35.206	As En Ma		178.272 8.306 45.113		179.057 8.133 43.881
As En Ma		188.477 7.677 37.78		187.826 7.06 28.649	As En Ma		177.5 8.037 47.437		175.051 7.913 48.414
As En Ma		188.128 7.62 34.951		188.374 7.571 34.488	As En Ma		177.845 8.123 41.481		179.262 8.086 44.646
As En Ma					As En Ma		180.804 8.134 38.554		182.596 8.229 42.547
As En Ma					As En Ma				176.906 7.995 45.003
As En Ma					As En Ma				182.127 8.095 36.98
As En Ma					As En Ma				184.687 8.041 32.267

Continuation of **Table 1**

5.		*Ṭā'*			6.		Final 'ayn		
	A		B			A		B	
	symbol	parameter	symbol	parameter		symbol	parameter	symbol	parameter
As En Ma		188.491 7.925 45.0		190.379 7.418 40.517			189.226 7.45 32.024		191.242 7.109 43.139
As En Ma		188.026 8.002 46.078		189.948 7.575 41.568			189.311 7.749 36.823		191.505 6.837 42.477
As En Ma		188.612 7.932 44.17		190.262 7.518 43.574			190.522 7.574 42.664		190.412 7.257 32.676
As En Ma		186.884 7.709 36.424		189.941 7.686 45.274			190.134 7.571 39.136		190.802 7.38 41.598
As En Ma		188.164 8.038 47.29		190.062 7.565 44.5			189.834 7.634 39.556		190.439 7.474 43.009

7.		Middle *mīm*			8.		Initial *jīm/ḥā'/khā'*		
	A		B			A		B	
	symbol	parameter	symbol	parameter		symbol	parameter	symbol	parameter
As En Ma		178.565 8.467 39.144		180.497 8.23 36.718			182.397 8.042 35.369		186.744 7.62 29.057
As En Ma		179.925 8.183 36.318		179.25 8.29 37.415			181.453 7.775 33.673		184.939 7.886 32.831
As En Ma		180.112 8.471 37.245		179.292 8.214 35.537			183.92 7.963 32.982		183.291 8.051 35.522
As En Ma		179.669 8.281 40.219		181.053 8.424 35.347			183.542 7.844 31.681		181.586 8.013 35.983

*Continuation of **Table 1***

7.	Middle *mīm*				8.	Initial *jīm/ḥā'/khā'*			
	A		B			A		B	
	symbol	parameter	symbol	parameter		symbol	parameter	symbol	parameter
As		179.909		177.953	As		182.589		185.282
En		8.367		8.452	En		7.922		7.738
Ma		42.024		40.455	Ma		35.128		29.545
As					As		182.934		182.61
En					En		7.947		8.143
Ma					Ma		35.461		36.764
As					As		185.22		184.411
En					En		7.867		7.888
Ma					Ma		31.719		32.062
As					As		185.259		
En					En		7.714		
Ma					Ma		30.85		

9.	Middle *hā'*				10.	*Lām/alif*			
	A		B			A		B	
	symbol	parameter	symbol	parameter		symbol	parameter	symbol	parameter
As		182.894		182.116	As		184.853		184.899
En		8.186		8.242	En		7.668		7.621
Ma		32.722		34.252	Ma		33.391		34.238
As		181.683		183.628	As		185.097		188.705
En		8.26		8.269	En		7.734		7.294
Ma		34.505		32.903	Ma		33.581		32.596
As		182.153		182.593	As		185.621		186.849
En		8.074		8.198	En		7.605		7.42
Ma		32.602		32.496	Ma		32.199		31.488
As		182.88		181.884	As		184.874		185.997
En		8.102		8.239	En		7.55		7.502
Ma		31.371		34.365	Ma		31.938		31.925
As		182.214		180.939	As		186.154		186.605
En		8.074		8.024	En		7.723		7.561
Ma		31.938		33.017	Ma		31.878		31.099

*Continuation of **Table 1***

11.					Final *qāf*				
	A		B		A		B		
	symbol	parameter	symbol	parameter	symbol	parameter	symbol	parameter	
As En Ma		188.538 6.999 28.849		189.591 7.202 27.071		189.777 7.183 28.578		189.41 7.136 26.031	
As En Ma		189.001 7.388 28.851		189.621 7.175 28.224					
As En Ma		188.567 7.153 29.863		188.71 7.028 27.966		189.744 6.911 29.152		186.507 7.326 27.346	

Table 2

Mean values of symbols

	Symbol	Hand	Number	As	En	Ma
1		A	5	185.954	7.3476	30.0050
	Initial *'ayn*	B	5	185.895	7.4326	30.9384
			10	*185.925*	*7.3901*	*30.4717*
2		A	5	189.303	7.1222	33.2352
	Alif	B	5	189.303	7.0080	33.4504
		All	10	*189.303*	*7.0651*	*33.3428*
3		A	5	188.270	7.5874	35.5306
	Final *nūn*	B	5	188.465	7.2658	32.5176
		All	10	*188.368*	*7.4266*	*34.0241*
4		A	6	180.021	8.1310	40.4802
	Middle *'ayn*	B	9	179.741	8.0597	41.7851
		All	15	*179.881*	*8.0953*	*41.1326*
5		A	5	188.035	7.9212	43.7924
	Ṭā'	B	5	190.118	7.5524	43.0866
		All	10	*189.077*	*7.7368*	*43.4395*
6		A	5	189.805	7.5956	38.0406
	Final *'ayn*	B	5	190.880	7.2114	40.5798
		All	10	*190.343*	*7.4035*	*39.3102*
7		A	5	179.636	8.3538	38.9900
	Middle *mīm*	B	5	179.609	8.3220	37.0944
		All	10	*179.623*	*8.3379*	*38.0422*

*Continuation of **Table 2***

Symbol		Hand	Number	As	En	Ma
8	Initial *jīm/ḥā'/khā'*	A	8	183.414	7.8843	33.3579
		B	7	184.123	7.9056	33.1091
		All	15	*183.769*	*7.8949*	*33.2335*
9	Middle *ḥā'*	A	5	182.365	8.1392	32.6276
		B	5	182.232	8.1944	33.4066
		All	10	*182.298*	*8.1668*	*33.0171*
10	*Lām/alif*	A	5	185.320	7.6560	32.5974
		B	5	186.611	7.4796	32.2692
		All	10	*185.965*	*7.5678*	*32.4333*
11	Final *qāf*	A	5	189.125	7.1268	29.0586
		B	5	188.768	7.1734	27.3276
		All	10	*188.947*	*7.1501*	*28.1931*

Table 3

Results of analytical treatment

(for $\alpha = 0.05$, C = 5.32 and $\alpha = 0.05$, C = 11.26)

Symbol			As	En	Ma
3	Final *nūn*	$F_{1.8}$	0.31846	**5.9701**	**5.37534**
		$\alpha = 0.05$	—	+	+
		$\alpha = 0.01$	—	—	—
5	*Ṭā'*	$F_{1.8}$	**42.6509**	**26.3858**	0.11159
		$\alpha = 0.05$	+	+	—
		$\alpha = 0.01$	+	+	—
6	Final *'ayn*	$F_{1.8}$	**10.7735**	**9.92659**	0.9074
		$\alpha = 0.05$	+	+	—
		$\alpha = 0.01$	—	—	—
10	*Lām/alif*	$F_{1.8}$	3.70392	**6.93993**	0.2451
		$\alpha = 0.05$	—	+	—
		$\alpha = 0.01$	—	—	—
11	Final *qāf*	$F_{1.8}$	0.30403	0.23913	**15.24**
		$\alpha = 0.05$	—	—	+
		$\alpha = 0.01$	—	—	+

Notes

1. E. A. Rezvan, N. S. Kondybaev, "New tool for analysis of handwritten script", *Manuscripta Orientalia*, II/3 (1996), pp. 43—53.

2. E. A. Rezvan, "The Qur'ān and its world: VI. Emergence of the Canon: the struggle for uniformity", *Manuscripta Orientalia*, IV/2 (1998), pp. 13—54.

3. F. Déroche, *The Abbasid Tradition. Qur'āns of the 8th to the 10th Centuries AD.* (Oxford, 1992), pp. 11—2. — The Nasser D. Khalili Collection of Islamic Art, II.

4. E. Whellan, "Writing the word of god: some early Qur'ān manuscripts and their milieu", pt. I, *Ars Orientalis*, XX (1990), pp. 119—23, figs. 19—22.