

トルファンマニ文字コードとグリフのデザイン

Design of Turfan Manichaean Character Code and Glyphs

ウメルジアン ウスマン†

Omarjan Osman

中平 勝子†

Katsuko T. Nakahira

三上 喜貴†

Yoshiki Mikami

山田 耕一†

Yamada Koichi

概要

トルファンは、中央アジアの天山山脈の東南にあった古代国名であり、トルファン盆地はトルファン国の一部であった。トルファンウイグル王国と呼ばれる。トルファンウイグル王国の言語は古代チュルク語(英: Old Churuk Language)である。紀元1世紀から7世紀までにトルファンウイグル王国で国字として使われていた文字は、トルファン文字と呼ばれている。マニ教(Manichaeism)の貴重な資料がトルファン文字で記録しているため、トルファンマニ文字とも呼ばれており、ウイグル族以外にウズベク(Uzbek)、タタール(Tatar)、キルギス(Kirghiz)、カザフ(Kazakh)、サラル(Salar)、ノガイ(Nogai)語系の民族もトルファンマニ文字を使った記録事がある。トルファンマニ文字は8個の大母音字、8個の小母音字、24個の大字音字、24個の小字音字、2個の分音記号、7個の句読点記号、10個の数字から構成されている。トルファンマニ文字で書かれた文献作品は、現在各国の博物館と大学に所蔵されている。しかしながら、トルファンマニ文字コード表は国際標準であるISO/IEC 10646及びUnicodeに含まれておらず、貴重な古代文献資料の電子的手段による流通の障害となっている。文字名に関する国際標準ISO/IEC 15942にもトルファン文字の名前は登録されていない。マニ文字コード表がISO/IEC JTC1/SC2に提案されているが[4]、その文字コード表にトルファンマニ文字は含まれていない。

本研究では、ドイツの調査隊が20世紀初頭にトルファンで発見した資料から字形を抽出し、文字コード国際規格の基本的な原則であるコード=グリフ分離原則[5]に従って、トルファンマニ文字の文字コード表とグリフ表のデザインを行った。その現状を報告する。

キーワード: トルファンマニ文字、文字コード、グリフ

1 はじめに

1.1 文字の起源と系統

マニ文字はアラム文字から発生した文字であり、紀元前1世紀頃にイランで使われはじめた。イランで誕生したマニ文字は西はイランから東はトルファンに至る広範囲で使われるようになったが、利用者達の言語・音韻体系が異なることから文字体系としても相違を生じており、筆者は、イランマニ文字、トルファンマニ文字、ソグドマニ文字の3つを区別すべき異なる文字体系と考える。このうち、本報告の主題であるトルファンマニ文字は紀元後1世紀にトルファンで使われた文字である。トルファンマニ文字を基礎として紀元後5世紀頃にソグディアナ地方でソグド語の利用者達によってソグドマニ文字が作成された(図1参照)。

紀元7世紀にトルファンウイグル王国がウイグル文字を国字として使った。トルファンマニ文字もトルファン人の一部が第二国字として使った記録事がある。紀元後12世紀にモンゴル王国とトルファンウイグル王国ではウイグル文字が国字として広く使われた。紀元後15世紀にモンゴル人とトルファン人の一部がイスラム教の影響を受けてアラビア文字を使った。紀元後16世紀からトルファンマニ文字を使う人は非常に少なくなっていった。トルファンマニ文字列は、母音字、子音字、分音記号と複雑なりガチャーの組み合わせから構成されている。それらの書き方は右から左の横書きであり、上から下の縦書きは存在しない。

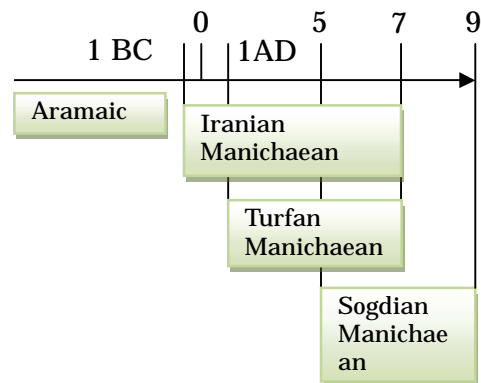


図1 古代から現代までの文字の変化

1.2 文字の分類

トルファンマニ文字は大文字(Capital)と小文字(Small)のケースを持っている。通常は小文字を用いるが、文書の書き起こしに大文字を用いページの初めには大文字とも区別される独特の書体で表現する。本稿では、この字形を「サインシンボル」と名付けた。本の表紙、文章のタイトルは大文字で表現する。この使用法はラテン文字やキリル文字のケース使用法にも類似していることから、コード化にあたっては、一つの文字に対して、大文字、小文字の二つのケースに別の符号位置を割り当てるといった考えをとった。

例: 以下の文の意味は、(私は本を読む)。大文字(𐰪), 小文字(𐰫), サインシンボル(𐰬)で表現する。

1) 文書の初めの第1文字は大文字(𐰪)で始まって、次の第2文字はアルファベットのの小文字を使う事がある。

𐰪𐰫𐰬𐰭𐰮𐰯𐰰𐰱𐰲𐰳𐰴𐰵𐰶𐰷𐰸𐰹𐰺𐰻𐰼𐰽𐰾𐰿

2) 1ページの初めサインシンボル(𐰬)で始まって、次の第2文字はアルファベットのの小文字を使う事がある。第2文字には大文字を使わない。

† 長岡技術科学大学
新潟県長岡市上富岡町 1603-1

Ⲁ ⲁ Ⲃ ⲃ Ⲅ ⲅ Ⲇ ⲇ Ⲉ ⲉ Ⲋ ⲋ Ⲍ ⲍ Ⲏ ⲏ Ⲑ ⲑ Ⲓ ⲓ Ⲕ ⲕ Ⲗ ⲗ Ⲙ ⲙ Ⲛ ⲛ Ⲝ ⲝ Ⲟ ⲟ Ⲡ ⲡ Ⲣ ⲣ Ⲥ ⲥ Ⲧ ⲧ Ⲩ ⲩ Ⲫ ⲫ Ⲭ ⲭ Ⲯ ⲯ Ⲱ ⲱ Ⲳ ⲳ Ⲵ ⲵ Ⲷ ⲷ Ⲹ ⲹ Ⲻ ⲻ Ⲽ ⲽ Ⲿ ⲿ Ⲁ ⲁ Ⲃ ⲃ Ⲅ ⲅ Ⲇ ⲇ Ⲉ ⲉ Ⲋ ⲋ Ⲍ ⲍ Ⲏ ⲏ Ⲑ ⲑ Ⲓ ⲓ Ⲕ ⲕ Ⲗ ⲗ Ⲙ ⲙ Ⲛ ⲛ Ⲝ ⲝ Ⲟ ⲟ Ⲡ ⲡ Ⲣ ⲣ Ⲥ ⲥ Ⲧ ⲧ Ⲩ ⲩ Ⲫ ⲫ Ⲭ ⲭ Ⲯ ⲯ Ⲱ ⲱ Ⲳ ⲳ Ⲵ ⲵ Ⲷ ⲷ Ⲹ ⲹ Ⲻ ⲻ Ⲽ ⲽ Ⲿ ⲿ

トルファンマニ文字と文字コード表の定義と略称を以下のように示す。

母音文字を“Vowel(V)”、子音文字を“Consonant(C)”、分音記符号を“Diacritics Marks(D)”、句読点記号を“Punctuation Marks(P)”、数字を“Digit(I)”、サインシンボルを“Signature Symbol(S)”で表現する。

本稿では、トルファンマニ文字を8大母音字、8小母音字、24個の大字音字、24個の小字音字、二つの分音記符号、7個の句読点記号、10個の数字、8個の大母音字サインシンボル、8個の小母音字サインシンボル、24個の大字音字サインシンボル、24個の小字音字サインシンボルに分類した。(表1参照)。

表1トルファン文字の分類

Acronym	Name
V	Capital Vowels
	Small Vowels
C	Capital Consonants
	Small consonants
D	Diacritics Marks
P	Punctuation Marks
I	Digits
S	Capital Vowels Signature Symbol
	Capital Consonants Signature Symbol
	Small Vowels Signature Symbol
	Small Consonants Signature Symbol

1.3 文字集合とコード表提案

トルファンマニ文字は母音と子音を表す表音文字からなる文字体系であり、文字集合は、a、æ、e、i、o、ø、u、yとの各母音に対応する8種の大母音字と小母音字、b、p、t、ʄ、ʃ、x、d、r、z、s、ʁ、f、q、k、g、ŋ、l、m、n、h、v、jの各子音に対応する24種の大子音字と小子音字がある。トルファンマニ文字はチュルク語系民族が使った主要なアルファベットのひとつである。表2に各文字の大文字と小文字の字形(いずれも独立形)、対応する音価、及び文字コード表上の相対位置を示す。相対位置はX000を起点とした16進数で示した。Xの値は、今後国際標準化のプロセスで決定されるものである。また、2段目の()内に示した数字は表3との対応を示したものである。

表2トルファンマニ文字の文字集合

Ⲁ ⲁ	Ⲃ ⲃ	Ⲅ ⲅ	Ⲇ ⲇ
I /i/ (1) X023, X003	E /e/ (3) X022, X002	Ah /æ/ (2) X021, X001	A /a/ (1) X020, X000
Uv /y/ (8) X027, X007	U /u/ (7) X026, X006	Ov /ø/ (6) X025, X005	O /o/ (5) X024, X004
Zhe/ʄ / (12) X02B, X00B	Te /t/ (11) X02A, X00A	Pe /p/ (10) X029, X009	Be /b/ (9) X028, X008
Re /r/ (16) X02F, X00F	De /d/ (15) X02E, X00E	He /x/ (14) X02D, X00D	Che /ʃ / (13) X02C, X00C

Ⲉ ⲉ	Ⲇ ⲇ	Ⲅ ⲅ	Ⲃ ⲃ
Shi / / (20) X033, X013	Si /s/ (19) X032, X012	Zhee / / (18) X031, X011	Ze /z/ (17) X030, X010
Ke /k/ (24) X037, X017	Khi /q/ (23) X036, X016	Fi /f/ (22) X035, X015	Ghe /ʁ/ (21) X034, X014
Me /m/ (28) X03B, X01B	Le /l/ (27) X03A, X01A	Ngi /ŋ/ (26) X039, X019	Ge /g/ (25) X038, X018
Ye /j/ (32) X03F, X01F	Ve /v/ (31) X03E, X01E	Hee /h/ (30) X03D, X01D	Ne /n/ (29) X03C, X01C

1.4 グリフの分類

マニ文字は、その先祖であるアラム文字をはじめとするセム系文字の性質を継承しており、同一文字であっても語中の位置によってその表示形を変える。すべての母音文字、子音文字に、語頭形、語中形、語末形、独立形の4つの表示形がある。これらの異なる表示形のそれぞれに別の符号位置を割り当てるという方法もあるが、文字コードに関する国際規格の基本原則は「コードとグリフの分離」[5]を定めているため、コード表とは別に、各表示形に対応するグリフ(Glyph、字形を示す抽象図形)をデザインした。

4種類のグリフの一般的な表現をG_iとし、個々の表現は、語頭形をG_l(Left-joining glyph form)、語中形をG_{cm}(Dual-joining glyph form)、語末形をG_r(Right-joining glyph form)、独立形をG_n(Nominal glyph form)とする。文字を単母音と単子音に区別する必要がある場合、一般形はG_iの代わりに、それぞれG_l、G_rを使用し、4種類の語頭形、語中形、語末形、独立形のグリフは、それぞれG_{vn}、G_{vr}、G_{vm}、G_{vl}(単母音の場合)、G_{cn}、G_{cr}、G_{cm}、G_{cl}(単子音の場合)のように表す。

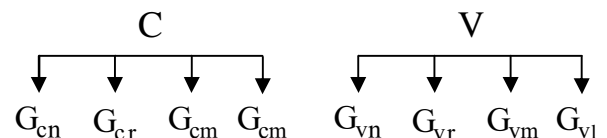


図2トルファングリフの実装

独立形大グリフと独立形小グリフと二つタイプが存在している。独立形大グリフを“Capital Glyph”、独立形小グリフを“Small Glyph”で表現する(表3参照)。

表3トルファンマニ文字のグリフ一覧

No	Turfan Name	G _n		G _r	G _{cm}	G _l	音価
		Capital	Small				
1	A	Ⲇ	ⲇ	Ⲉ	ⲉ	Ⲋ	/a/
2	AH	Ⲇ	ⲇ	Ⲉ	ⲉ	Ⲋ	/æ/
3	E	Ⲃ	ⲃ	Ⲅ	ⲅ	Ⲇ	/e/
4	I	Ⲁ	ⲁ	Ⲃ	ⲃ	Ⲅ	/i/
5	O	Ⲇ	ⲇ	Ⲉ	ⲉ	Ⲋ	/o/
6	OV	Ⲃ	ⲃ	Ⲅ	ⲅ	Ⲇ	/ø/
7	U	Ⲃ	ⲃ	Ⲅ	ⲅ	Ⲇ	/u/
8	UV	Ⲃ	ⲃ	Ⲅ	ⲅ	Ⲇ	/y/

9	BE	𐰇	𐰇	𐰇	𐰇	𐰇	/b/
10	PE	𐰃	𐰃	𐰃	𐰃	𐰃	/p/
11	TE	𐰆	𐰆				/t/
12	ZHE	𐰑	𐰑	𐰑			/dʒ/
13	CHE	𐰏	𐰏	𐰏	𐰏	𐰏	/tʃ/
14	HE	𐰇	𐰇	𐰇	𐰇	𐰇	/x/
15	DE	𐰃	𐰃	𐰃			/d/
16	RE	𐰃	𐰃	𐰃			/r/
17	ZE	𐰑	𐰑	𐰑			/z/
18	ZHEE	𐰑	𐰑	𐰑			/ /
19	SI	𐰏	𐰏	𐰏	𐰏	𐰏	/s/
20	SHI	𐰏	𐰏				/ /
21	GHE	𐰑	𐰑	𐰑	𐰑	𐰑	/ʒ/
22	FI	𐰏	𐰏	𐰏	𐰏	𐰏	/f/
23	KHI	𐰏	𐰏	𐰏	𐰏	𐰏	/q/
24	KE	𐰏	𐰏	𐰏			/k/
25	GE	𐰏	𐰏	𐰏	𐰏	𐰏	/g/
26	NGI	𐰏	𐰏	𐰏	𐰏	𐰏	/ /
27	LE	𐰏	𐰏	𐰏	𐰏	𐰏	/l/
28	ME	𐰏	𐰏	𐰏	𐰏	𐰏	/m/
29	NE	𐰏	𐰏	𐰏	𐰏	𐰏	/n/
30	HHE	𐰏	𐰏	𐰏	𐰏	𐰏	/h/
31	VE	𐰏	𐰏	𐰏	𐰏	𐰏	/v/
32	YE	𐰏	𐰏	𐰏	𐰏	𐰏	/j/

1.5 トルファンリガチャー

連続するグリフの組み合わせが単純な合成に従わず、独特の合成字形を形成する場合がある。これをリガチャー(Ligature)と呼ぶ。リガチャーの集合をL={L_n, L_r, L_m, L_l}で表現する。リガチャーもまたグリフ表に追加する必要がある。

トルファンマニ文字のリガチャーには以下の5通りの組み合わせがある。

- 1) 語頭形グリフG₁+語中形グリフG_mの組み合わせ
G₁+G_m=L_l (語頭形: Left-Joining ligature)
- 2) 語頭形グリフG₁+語末形グリフG_rの組み合わせ
G₁+G_r=L_n (独立形: Nominal ligature)
- 3) 語中形グリフG_m+語中形グリフG_mの組み合わせ
G_m+G_m=L_m (語中形: Dual-joining ligature)
- 4) 語中形グリフG_m+語末形グリフG_rの組み合わせ
G_m+G_r=L_r (語末形: Right-joining ligature)
- 5) 語頭形グリフG₁+語中形グリフG_m+語末形グリフG_rの組み合わせ
G₁+G_m+G_r=L_n (独立形: Nominal ligature)

表4 トルファンマニ文字リガチャーの形成

グリフ(G _n)	グリフ(G _n)	リガチャー(L _n)
𐰏 + 𐰏	𐰏 = 𐰏	𐰏𐰏

リガチャーの二重母音と二重子音グリフと集合をL̄={L̄_n, L̄_c}で表現する。

G_i+G_j+G_i=L̄_n, G_i+G_j+G_j=L̄_c

2 トルファン語とトルファンマニ文字の名称

トルファン語(古代チュルク語の一部)とトルファン文字(トルファンマニ文字)はの名称は、依然として言語名や文字名に関する国際規格に含まれていない。古代文字としてマニ文字の名前(英: Manichaean)は ISO 15924 Script code に登録されている。しかし、マニ文字は別々の三つの(イラン、トルファン、ソグド)マニ文字として区別されるべきと考えており、それぞれ独立に登録が必要である。言語名コードに関する国際標準 ISO 639-1 と ISO 639-2 にもトルファン語の名前は登録されていない。

表5 関連国際規格における選択された文字

Script Name	ISO 15924 Script code	ISO/IEC 10646 Unicode
Iran Manichaean	not present	not present
Turfan Manichaean	not present	not present
Sogdian Manichaean	not present	not present
Manichaean	Mani	not present

表6 選択された言語の ISO 630 言語コード

Script Name	ISO 639-1	ISO 639-2
Iran	---	ira
Iran Manichaean	not present	not present
Turfan	not present	not present
Turfan Manichaean	not present	not present
Sogdian	---	sog
Sogdian Manichaean	not present	not present

3 トルファンマニ文字フォントのデザイン

3.1 トルファンフォントのデザイン

Font Creator 5.6 を使って True Type フォントのデザインを以下のように行った。字形の典拠として用いたのは、1902年11月から1903年3月に、トルファン盆地でドイツの考古学者と遠征グループ“German Turfan expedition”が、発見した写本資料である。ちなみに、東トルキスタン“East Turkistan”の東ウイグル可汗国“ウイグル語: Sherqiy Uyghur Kagan Dowliti”の中期に記録されたマニ文字の多くの資料が、天山山脈のヤルクンタク“Yalkun tag(火炎の山)”の中にあったベゼクリク千仏洞“Bezekliik Ming Oy(装飾した千仏の洞)”で発見された。

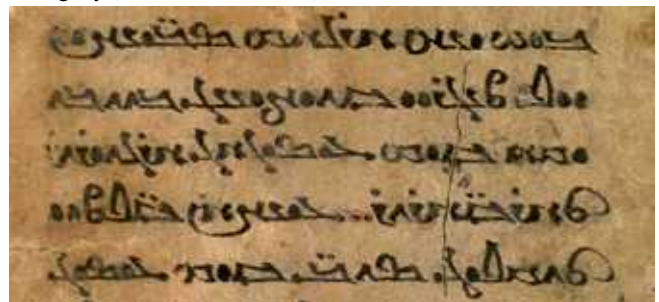


図3 トルファンマニ文字文献作品 (年代不明) u0074verso [1]

筆者は、トルファンマニ文字のデジタル処理の研究の参考文献として、現在ベルリントルファンコレクション“Berlin Turfan Collection [1]”に所蔵され、1904年に発表

されたトルファンマニ文字の写本資料 (u0007recto から u0161seite2、及び m0120_seite1 から m0834_seite1) を使用し、写本から各文字の字形を抽出し、これに基づいてフォントのデザインを行った。

また、フォント作成にあたって使用したツールおよび環境は以下の通りである。最終的には、トルファンマニ文字 True Type フォントとして 282 種のグリフを作成した。

1) New Font

- (1) Font family name: Turfan.
- (2) Character set: Unicode.
- (3) Font style: Regular.

2) Font file Type

True Type font (ttf)

3) Sort Glyphs

Sort by: Microsoft Unicode or Symbol code points.

4) Format Platform Manager

Platform: Microsoft Unicode BMP only.

5) Font Settings

Font direction hint: Only strong right to left.

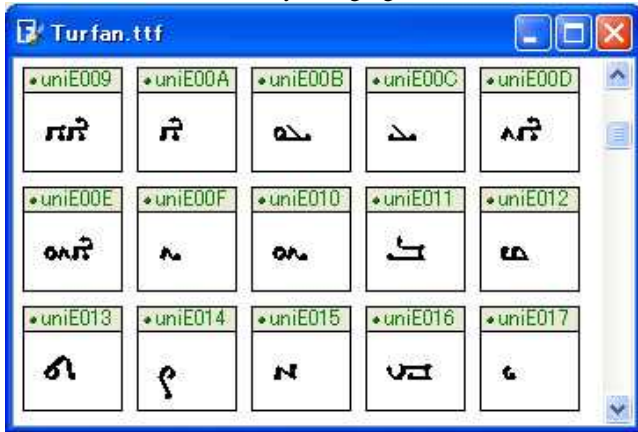


図4 トルファンフォント

3.2 トルファンマニ文字コード表

このようにして開発したフォントを用いて作成したトルファンマニ文字コードを表7に示す。トルファンマニ文字コード表には、8 大母音字(00-07)、8 小母音字(10-17)、24 個の大字音字(08-1F)、24 個の小字音字(20-3F)、二つの分音記号(40-41)、7 個の句読点記号(50-58)、10 個の数字(60-69)、8 個の大母音字サインシンボル(70-77)、8 個の小母音字サインシンボル(90-97)、24 個の大字音字サインシンボル(78-8F)、24 個の小字音字サインシンボル(98-AF)と、合計 147 種の文字のデザインを行った。

表7 トルファン文字コード表

		上位4ビット										
		0	1	2	3	4	5	6	7	8	9	A
下 位 4 ビ ット	0	𐰀	𐰁	𐰂	𐰃	𐰄	𐰅	𐰆	𐰇	𐰈	𐰉	𐰊
	1	𐰋	𐰌	𐰍	𐰎	𐰏	𐰐	𐰑	𐰒	𐰓	𐰔	𐰕
	2	𐰖	𐰗	𐰘	𐰙	𐰚	𐰛	𐰜	𐰝	𐰞	𐰟	𐰠
	3	𐰡	𐰢	𐰣	𐰤	𐰥	𐰦	𐰧	𐰨	𐰩	𐰪	𐰫
	4	𐰬	𐰭	𐰮	𐰯	𐰰	𐰱	𐰲	𐰳	𐰴	𐰵	𐰶
	5	𐰷	𐰸	𐰹	𐰺	𐰻	𐰼	𐰽	𐰾	𐰿	𐱀	𐱁
	6	𐱂	𐱃	𐱄	𐱅	𐱆	𐱇	𐱈	𐱉	𐱊	𐱋	𐱌
	7	𐱍	𐱎	𐱏	𐱐	𐱑	𐱒	𐱓	𐱔	𐱕	𐱖	𐱗
	8	𐱘	𐱙	𐱚	𐱛	𐱜	𐱝	𐱞	𐱟	𐱠	𐱡	𐱢
	9	𐱣	𐱤	𐱥	𐱦	𐱧	𐱨	𐱩	𐱪	𐱫	𐱬	𐱭

A	𐰁	𐰂	𐰃	𐰄				𐰈	𐰉	𐰊	𐰋
B	𐰌	𐰍	𐰎	𐰏				𐰒	𐰓	𐰔	𐰕
C	𐰖	𐰗	𐰘	𐰙				𐰜	𐰝	𐰞	𐰟
D	𐰡	𐰢	𐰣	𐰤				𐰨	𐰩	𐰪	𐰫
E	𐰬	𐰭	𐰮	𐰯				𐰴	𐰵	𐰶	𐰷
F	𐰺	𐰻	𐰼	𐰽				𐰿	𐱀	𐱁	𐱂

3.3 トルファンマニグリフ表のデザイン

トルファンマニ文字グリフの 29 種の母音グリフと 86 種の子音グリフのデザインを行った。

表8 トルファンマニ文字グリフ表

	000	001	002	003	004	005	006	007	008
0	𐰀	𐰁	𐰂	𐰃	𐰄	𐰅	𐰆	𐰇	𐰈
1	𐰉	𐰊	𐰋	𐰌	𐰍	𐰎	𐰏	𐰐	𐰑
2	𐰒	𐰓	𐰔	𐰕	𐰖	𐰗	𐰘	𐰙	𐰚
3	𐰛	𐰜	𐰝	𐰞	𐰟	𐰠	𐰡	𐰢	𐰣
4	𐰤	𐰥	𐰦	𐰧	𐰨	𐰩	𐰪	𐰫	𐰬
5	𐰭	𐰮	𐰯	𐰰	𐰱	𐰲	𐰳	𐰴	𐰵
6	𐰶	𐰷	𐰸	𐰹	𐰺	𐰻	𐰼	𐰽	𐰾
7	𐰿	𐱀	𐱁	𐱂	𐱃	𐱄	𐱅	𐱆	𐱇
8	𐱈	𐱉	𐱊	𐱋	𐱌	𐱍	𐱎	𐱏	𐱐
9	𐱑	𐱒	𐱓	𐱔	𐱕	𐱖	𐱗	𐱘	𐱙
A	𐱚	𐱛	𐱜	𐱝	𐱞	𐱟	𐱠	𐱡	𐱢
B	𐱣	𐱤	𐱥	𐱦	𐱧	𐱨	𐱩	𐱪	𐱫
C	𐱬	𐱭	𐱮	𐱯	𐱰	𐱱	𐱲	𐱳	𐱴
D	𐱵	𐱶	𐱷	𐱸	𐱹	𐱺	𐱻	𐱼	𐱽
E	𐱾	𐱿	𐲀	𐲁	𐲂	𐲃	𐲄	𐲅	𐲆
F	𐲇	𐲈	𐲉	𐲊	𐲋	𐲌	𐲍	𐲎	𐲏

4 まとめと今後の課題

写本資料に含まれるトルファンマニ文字写本資料に基づいて歴史的な文字の字形を調査し

Adobe Photoshop CS5 Extended、Font Creator 5.6 を用いて、文字コード表のデザイン、グリフのデザイン、True Type フォントのデザインを行った。現在、このトルファンマニ文字コード表について ISO への提案の準備と国際文字符号表の設計思想との整合性確認 (ISO 専門家) を行っている。

本研究を通じてトルファンマニ文字は大文字と小文字の字形を持っていること事がわかった。

今後の課題として入出力システムの試作と評価を進めてゆく予定である。

参考文献

[1] Turfanforschung of manuscripts from the Berlin Turfan-Collection, Digitales Turfan-Archiv.
 [2] The Worlds Writing Systems, edited by P. Daniels, W. Bright, New York, Oxford University Press, 1996.
 [3] 三上喜貴、文字符号の歴史：アジア編、2002 年
 [4] ISO/IEC JTC1/SC2/WG2 N4029R, L2/11-123R, 2011-05-10.
 [5] ISO/IEC TR 15285 An operational model for characters and glyphs, First edition. 1998-12-15.