

Proposal to encode ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE

Takehiko Kyono (京野長彦)

2026-10-17

1. Introduction

This proposal requests the encoding of one Arabic-script character used in the *yaña imlâ* (“New Orthography”; Tatar: яңа имля) writing systems of Tatar (Cyr. татар теле, Arb. تاتار تلي) and Bashkir (Cyr. башкорт теле, Arb. باشقورت تلي), in the *yañi imla* (Nogai: яңы имла) writing system of Nogai (Cyr. ногай тили, Arb. نوعاي ءتلي/نوعاي تلي), and in modern publications on Tatar language published in China. The character is based on ARABIC LETTER ALEF MAKSURA with a comma above. The possibility that the character was also used in Karakalpak (Lat. qaraqalpaq tili, Arb. قاراغالفاق ءتلي) was examined but not confirmed.

The character was previously proposed in L2/11-138 [1] under the name ARABIC LETTER FARSI YEH WITH DAMMA ABOVE, but was not encoded. This proposal reconsiders the encoding of that character in light of the current schematic analysis in ArabicShaping.txt and presents additional evidence of the character’s use.

2. Proposed character

2.1 Character information

Proposed Unicode Name:

ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE

Proposed Code Point: U+0892

Isolated	Final	Medial	Initial
ﺀ	ﺀ	ﺀ	ﺀ

Character Properties:

0892;ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE;Lo;0;AL;;;;;N;;;;;

ArabicShaping.txt:

0892; DOTLESS YEH WITH COMMA ABOVE; D; YEH

Its joining behavior is identical to that of U+0649 ARABIC LETTER ALEF MAKSURA.

2.2 Phonetic value

In historical Tatar and Bashkir Arabic-script orthographies in Tatarstan and Bashkortostan, the character represented a vowel whose phonetic realization depended on vowel harmony. In Tatar usage during the Soviet period, it represented /ʊ/ (modern Cyrillic: ы) in back-vowel contexts and /e/ (modern Cyrillic: е) in front-vowel contexts (Fig. 2, 3).

Examples:

Back-vowel: سورى (modern Cyrillic: соры)

Front-vowel: تۇسلى (modern Cyrillic: төсле)

In modern Tatar publications in China, the character represents /ʊ/ (modern Cyrillic: ы) consistently (Fig. 17(b)).

Example:

تاتارلارى (modern Cyrillic: татарлары)

In the historical Nogai Arabic-script orthography of 1927–1928, based on the author's analysis of the texts in [18], [19], [20], [21], [22], [23], the character represented /ʊ/ (modern Cyrillic: ы) in back-vowel contexts and /i/ (modern Cyrillic: и) in front-vowel contexts.

Examples:

Back-vowel: تۇرمۇس (modern Cyrillic: турмысын)

Front-vowel: ئىگىشى (modern Cyrillic: эгинши)

3. Background

3.1 Historical use in Tatar, Bashkir, and Nogai orthographies

The proposed character was used in the *yaña imlâ* (“New Orthography”; Tatar: яңа имля) writing systems of Tatar and Bashkir in the 1920s. It occurred as a distinct alphabetic character in these orthographic traditions.

The character is attested in materials documenting the Arabic-script period of Tatar and Bashkir writing (Fig. 1–10). L2/11-138 described the character among additional Arabic-script characters used for Tatar and Bashkir, and presented examples illustrating its use in early twentieth-century orthographic materials [1].

The proposed character was also used in the *yañı imla* (Nogai: янъы имла) orthography of Nogai [13, p. 166], which was introduced in 1926 and revised in 1927 [14]. The book in which the 1926 orthography was first published [15] uses an orthography resembling the Kazakh Arabic-script orthography known as *töte jazu* (Kazakh: төте жазу). The same is true of some books published in 1927 [16], [17]. None of these publications contains the proposed character. The character is attested in other books published in 1927 and in books published in 1928 (Fig. 11–16), which use an orthography resembling the Tatar *yaña imlâ*. Publication in Nogai in Arabic script ceased after 1928 [13, p. 166].

3.2 Continued use in modern Tatar publications in China

In addition to its historical use, the proposed character is attested in modern publications relating to the Tatar language in China.

Examples include a Tatar alphabet chart, a Tatar textbook, and a book publication containing the character in actual use (Fig. 17–19). These materials indicate that the character is not solely of historical interest, but remains relevant for the representation of Tatar written materials in China.

As the Tatar population is the smallest among China’s ethnic minorities, there are few materials in the Tatar language in China; however, the character proposed here is used in the aforementioned materials.

3.3 Previous proposal and encoding decision

The character was previously proposed for encoding in L2/11-138 under the name ARABIC LETTER FARSI YEH WITH DAMMA ABOVE [1]. During subsequent review, L2/11-201 considered this character representable by the sequence <U+0649 ARABIC LETTER ALEF MAKSURA, U+064F ARABIC DAMMA> (أُ), and therefore did not encode it as a separate character [2].

The present proposal revisits this decision based on additional evidence of use and a re-evaluation of the character’s identity and shaping behavior.

3.4 Karakalpak

L2/11-138 identifies the proposed character as being used in Karakalpak and provides an image of a Karakalpak alphabet chart as evidence [1, Fig. 2]. The same alphabet chart has been identified among historical Karakalpak materials examined for this proposal (Fig. 20).

However, examination of three historical books in Arabic-script Karakalpak, including the book containing this alphabet chart, has not revealed the proposed character in running text (Fig. 21–23). The orthographic conventions found in these materials are instead close to those of the Kazakh *töte jazı*.

The presence of the character in an alphabet chart therefore does not appear sufficient to establish that it was actually used as a distinct character in Karakalpak. Accordingly, this proposal does not identify Karakalpak as a language in which the character was used.

4. Encoding rationale

4.1 Distinction from ALEF MAKSURA + DAMMA

The proposed character was previously considered representable by the sequence <U+0649 ARABIC LETTER ALEF MAKSURA, U+064F ARABIC DAMMA> in L2/11-201 [2].

However, this sequence does not accurately represent the identity or behavior of the character. In the historical orthographies where the character was used, the mark above is not an independent vowel sign applied to ALEF MAKSURA, but forms an integral part of the letter.

The proposed character is represented as a separate alphabetic character in the relevant orthographic traditions. Its representation as ALEF MAKSURA followed by DAMMA obscures this distinction.

4.2 Letter identity and joining behavior

The base shape of the proposed character corresponds to ARABIC LETTER ALEF MAKSURA rather than FARSI YEH. In particular, when used in connected text, the character does not acquire the two dots below that characterize FARSI YEH. Its joining behavior is therefore consistent with ALEF MAKSURA.

The proposed character should consequently be encoded as a letter based on ARABIC LETTER ALEF MAKSURA with a comma above, rather than as a sequence of an existing letter and a combining mark.

4.3 Analysis of the mark above the letter

The character was originally proposed in L2/11-138 as ARABIC LETTER FARSI YEH WITH DAMMA ABOVE [1]. However, the mark above the letter is not a conventional vowel sign indicating a short vowel, but an element forming part of the letter identity.

A similar case is found in the Unicode representation of ARABIC LETTER U (U+06C7). Earlier versions of ArabicShaping.txt described ARABIC LETTER U using a damma-based schematic name, while the corresponding entry in ArabicShaping.txt was revised in Unicode 14.0 to describe this character with a comma above [3, Sect. 11a, pp. 21–22]. (ﻯ → ﺀ)

This change reflects a more accurate description of the graphical relationship between the base letter and the mark above it. The proposed character should likewise be analyzed as an ALEF MAKSURA-based letter with a comma above, rather than as ALEF MAKSURA followed by ARABIC DAMMA.

[8] (М. Гафури), *Хазер ансат ул, элек кыен иде* (Хазер ансат ул, элек кыен иде): *ئلك قىين ئيدى*, م. عهفورى [8]

СССР (СССР) س.س.س.ر. حالقلارنىڭ ئوزەك نەشرىياتى: (Мәскәү), *Балалар өчен* (Балалар өчен) *لالار ئوچن* халыкларының үзәк нәшрияты), 1927.

[9] David Trilling, “Central Asia: Propaganda Show Spotlights Soviet Push in Muslim Lands,” *Eurasianet*, Sep. 22, 2013. [Online]. Available: <https://eurasianet.org/central-asia-propaganda-show-spotlights-soviet-push-in-muslim-lands>. [Accessed: Sep. 19, 2026].

[10] (В. Хангилдин), and (Н. Таһиров), *Н. Таһиров*, (Х. Ғабидов), *Х. Ғабидов*, *ع. عابيدف* [10] (Өфө), *Башкорт теленең сарфы* (Башкорт теленең сарфы) *ئۇفۇ* (Өфө), USSR: (Ф. Вилданов), *ع. ويلدانف* (Башкнига); *Наркомпрос* (Наркомпрос), 1925. <https://disk.yandex.ru/a/eBqV9hZG3V792R>.

[11] (Н. Таһир), *Башкорт алифбаһы* (Башкорт алифбаһы). Уфа, USSR: *Мәркәз шәрқ нәшрияты* (Мәркәз шәрқ нәшрияты) *«Башкнига» ширкәте*, «Башкнига» ширкәте; *Ширкәте* (Ширкәте), 1924.

[12] *Башкортостан* (Башкортостан), Jul. 1, 1927. <https://archive.org/details/bashqortostan--npp/Башкортостан.%20-%201927.%20-%20II%20полугодие./page/n309/mode/2up>.

[13] А. Х. Курмансеитова, *У истоков ногойской книги (XIX – начало XX века): монография*. Черкесск, Russia: Карачаево-Черкесский ордена «Знак Почёта» институт гуманитарных исследований (КЧИГИ), 2009. ISBN: 978-5-85183-018-1.

[14] А. Х. Курмансеитова, “О переходе ногойцев от арабской графики к латинице в годы нэпа,” *Исламоведение*, no. 2, pp. 46–51, 2011. [Online]. Available: <https://islam.dgu.ru/archive.aspx?idS=130>.

[15] (Ресей), *Ресей* *له نين شى كه مونيس جاسلارى بيرگى نىڭ جورىعن ئۇسندىرئ*, (С. Кемрад), *С. Кемрад*, *س. كه مرهد* [15] (Н. Ногайлы). *Ленинши кемунис яслары биригининъ йорыгын туъсиндирув*, trans. *نو عايلى* (Н. Ногайлы). СССР эллерининъ орта баспасы), *Мескув* (Мескув), USSR: *هله رىنلڭ ورتا باسباسى* (Мескув), USSR: *ميسكو* (Мескув), 1926.

[16] (Н. Ногайлы). *Ленин эм эгинишлер* (Ленин эм эгинишлер), trans. *نو عايلى* (Н. Ногайлы). СССР эллерининъ орта баспасы), *هله رىنلڭ ورتا باسباسى* (Мескув), USSR: *ميسكو* (Мескув), USSR: *ميسكو* (Мескув), 1927.

- [17] (Балаларга) بالالارغا (Биздин устатымыз Ленин): ۇستатىمىز لەنين, (З. Лилина), ز. لىلىنا [17], trans. (СССР) س.س.س.س.ر. ەللەرىنىڭ ورتا باسباسى (Мескуёв), USSR: (Н. Ногайлы) ن. نوغايلى, 1927. эллерининъ орта баспасы), 1927.
- [18] (Он йыллык куьрес эм туьзилис) ئون يىللىق كۈرۈس ئىم تۈزۈلۈش, (И. С. Новаковский) ئى. س. نەۋەكوپسكى [18], trans. (Мескуёв), USSR: (Н. Ногайлы) ن. نوغايلى, 1927. эллерининъ орта баспасы), 1927.
- [19] (8-нши март янъы аьмел): 8-نشى مارت جاъمى ئەمىل, (Н. Ногайлы) ن. نوغايلى [19], trans. (Мескуёв), USSR: (Хатын-кыз тенълиги туврасында) قاتىن-قىز تىڭلىڭى توراسىدا, 1927. эллерининъ орта баспасы), 1927.
- [20] (1-май байрамы) 1-ماى بايرامى, (И. Сорин) ى. سورىن [20], trans. (Мескуёв), USSR: (Н. Ногайлы) ن. نوغايلى, 1927. эллерининъ орта баспасы), 1927.
- [21] (Сифилис) سىفلىس تىرىپىر شەنكر: (Венера аврувлары) ۋىنېرو آۋرۇلارنى, (Доктор Броннер) دوكتور بروننېر [21], trans. (Мескуёв), USSR: (Н. Ногайлы) ن. نوغايلى, 1927. эллерининъ орта баспасы), 1927.
- [22] (Хатын-кызлар арасында саясий анъ-билим юртларынынъ ислери): مەتود قاتى (Метод каты), (Гали Ибраимов) ەلى ئىبراىمف [22], trans. (Мескуёв), USSR: (Н. Ногайлы) ن. نوغايلى, 1928. эллерининъ орта баспасы), 1928.
- [23] (Аьруьв (саният) ۋىرۇ (سەنەت) جىرلارنىنان. قارايدار مىنان قىز كۈل, (А. Яны Бек) ئ. جانى بەك [23], trans. (Мескуёв), USSR: (Н. Ногайлы) ن. نوغايلى, 1928. эллерининъ орта баспасы), 1928.
- [24] 热合甫·阿巴斯 and 周建华, 中国塔塔尔族 银川, China: 宁夏人民出版社, 2012. ISBN: 978-7-227-04775-9.
- [25] “图文-祝福北京塔塔尔族使者评选 记录塔塔尔的书籍,” 新浪奥运, Nov. 1, 2007. [Online]. Available: <https://2008.sina.com.cn/p/2007-11-01/140231682.shtml>. [Accessed: Jul. 29, 2026].

تاشكەن. (Egedeler sawatı) سېگەدەلەر ساۋاتى, (Seyfulgábit Májitov) سەيفىل غايىت مەجىتف [26]
(Ташкен), USSR: قازاقستان مەملەكەت باسباسىنىڭ كۇن شعس ءبولمى (Қазақстан мемлекет
баспасының күншығыс бөлімі), 1925.

تاشكەن. (Qaraqalpaq tilinde álipbe) قاراغالپاق تىلىندە الپبە, (S. Májitov) س. مەجىدوف [27]
(Ташкен), USSR: قازاقستان مەملەكەت باسباسىنىڭ كۇن شعس ءبولمى (Қазақстан мемлекет баспасының
күншығыс бөлімі), 1925.

تاشكەن. (Oqıw kitabı) وقۇر كىتابى, (S. Májitov) س. مەجىدوف [28]
(Ташкен), USSR: قازاقستان مەملەكەت باسباسىنىڭ كۇن شعس ءبولمى (Қазақстан мемлекет баспасының күншығыс
бөлімі), 1925.

6. Figures

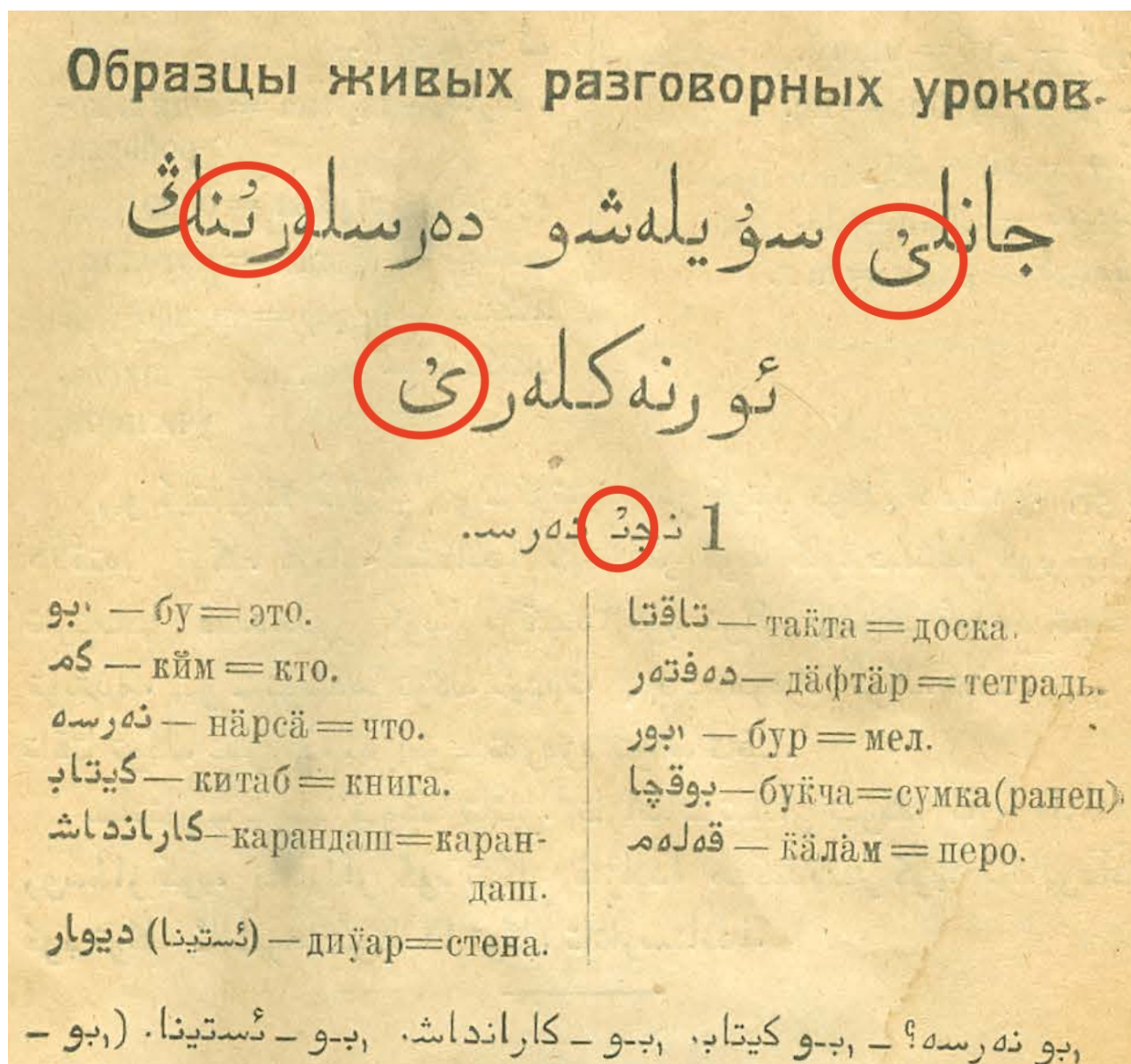


Fig. 1. Tatar language in [4]. It includes initial, medial, final, and isolated forms.

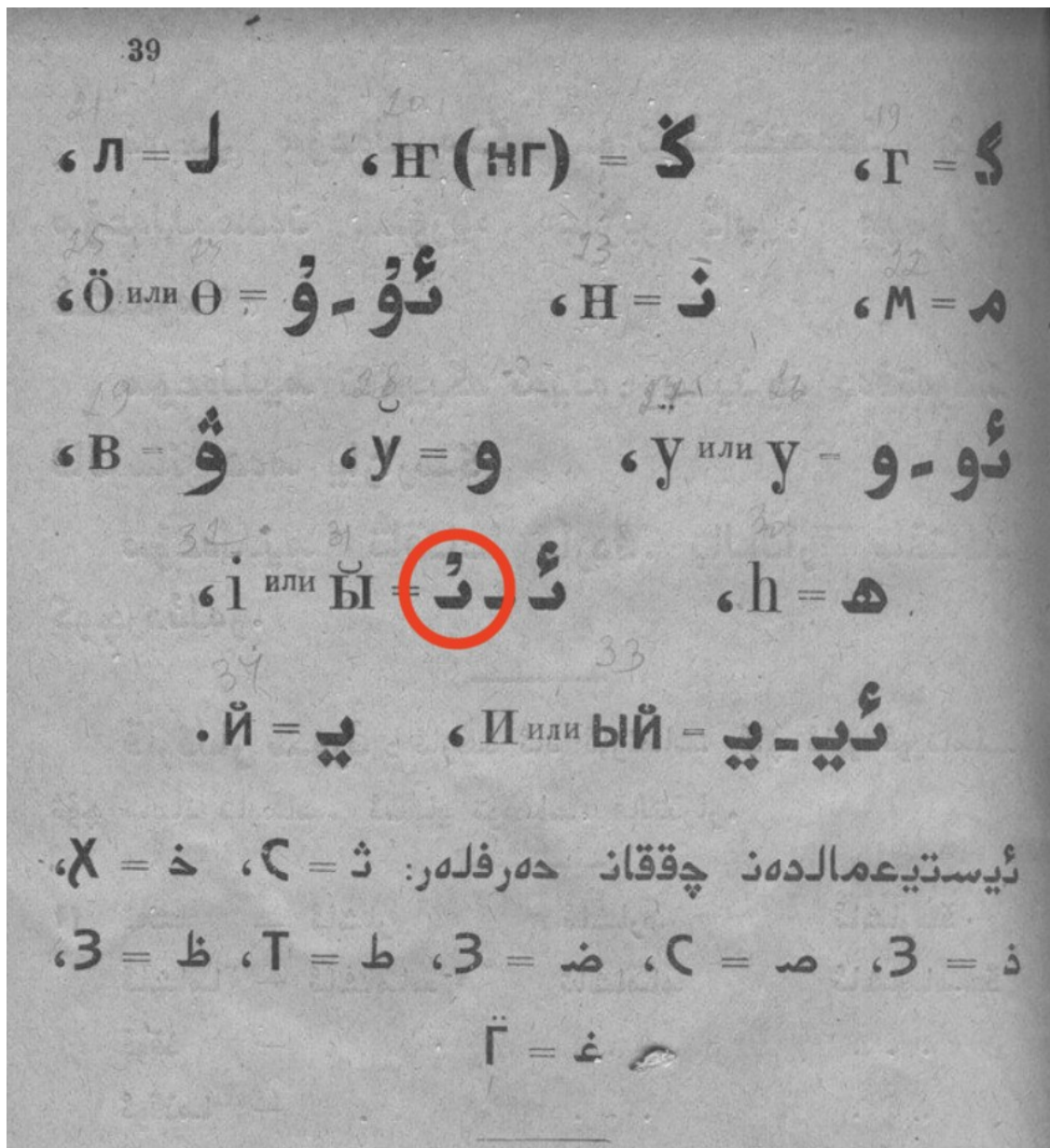


Fig. 2. Tatar alphabet in [5]. Note that the corresponding Cyrillic letters are i and Ы, since the Cyrillic orthography was not well-developed.

تەلىپبا تەرتىپى					
تەرتىپ	ھەرڧلەرنىڭ	ئورس تەلىپباسىنىڭ			
	نومىرلارنى	يازاڭلىقىنى			
1	ئ		С	س	13
2	ا		Ш	ش	14
3	ب		Г	غ	15
4	پ		Ф	ف	16
5	ت		К	ق	17
6	ج		К	ك	18
7	چ		Г	گ	19
8	ح		Н	ك	20
9	د		Л	ل	21
10	ر		М	م	22
11	ز		Н	ن	23
12	ژ		Ў, У	و	24
			Б	ق	25
			Ө, Ö	و	26
			h	ھ	27
			А	ه	28
			И, Й, Ы	ي	29
			И, Ы	ئ	30

Fig. 3. Tatar alphabet in [6]. Note that the corresponding Cyrillic letters are й and Ы, since the Cyrillic orthography was not well-developed.

ئىككىنچى باسماغا مۇقەددىمە.

«تاتار ئەدەبىياتى» تارىخىنىڭ ئىككىنچى ھەم ئۈچىنچى بۆلۈكلەرنىڭ ئىككىنچى باسماسىدا كىرەك تۈزۈلۈش، كىرەك قاراش ياعنىدىن ئەساسى ئۆزگەرتىشلەر ياسالدى. بۇ ئۆزگەرتىشلەر باشلىغا شۇنداقلا غەيرەت:

بۇرنى تاتار ئەدەبىياتىنىڭ بۇتنى ئەساسى قىسمى بىر بۇتنى ھالەتدە، شۇنى كىتابقا كىرگۈزۈپ بىرلىككە كەلتۈرۈلدى. تاتار ئاراسىدا چىقماق ۋە غۇسمانلىقلاردان كىر، ئۆزگەرتىلمىدى. ئۆزلىشىدىن كەلگەن ئەسەرلەر، ئاسل تاتار ئەدەبىياتىنىڭ بارلىقى ياعنىدىن ئەدەبىيەتلەرنى ئىككىنچى دەرىجەدە بولغانغا، تۈپ ماتېرىيالنىڭ بۇ تىللىشىغا زارار كىتمەس ئۈچۈن، ئايرىلىپ چىقارلدى.

بۇلار بىر تىل بولۇپ بىرلىككە كەلتۈرۈلۈپ ئايرىم بىر كىتاب بولۇپ چىقماق.

Fig. 4. Tatar language in [7]. It includes initial, medial, final, and isolated forms.

عەلى ئابزى

عەلى ئابزى 50 بەشكە جىنب بارا، ئانك چەچلەرى بارتىلاش
ئاعارغان. ئارقاسى بۇكرەبە باشلاغان. كوز تۇبلەرى جىبىرچقلاىب، بۇرۇلب
تۇرا. كوزلەرى ناچارلانا باشلاغانغا كورە، سۇگىى بللاردا كوزلك كىيە
باشلاغان. ئانى كورو بلەن، ئوز عۇمۇرنىدە بىك كوب ئشلەگەنلىكى
كورۇنب تۇرا. ئول شولاى قارتايا تۇشسەدە بىك شاد. ئىرتەنگى سەعت
توعزىدان ئالب، تۇشدەن سۇك سەعت دورتكە چاقلى ئانى ئش ئورنىدا
كۆرەسك. ئول شول جىدى سەعت ئش واقتىنىك برگنە مىنوتىدا بوشقا
ئوتكەرمى. ئانك ئشنى ئاورسۇنب، زارلاب سۇيلەگەنن ئىشتكەن كشى
بوق. ئول ئوزۇنىك ھەممە ئشلەرن پۇختە ھەم تەرتىب بلەن ئالب بارا.
ئوزۇنىك ئۇبىندەدە كوب واقتىن برەر ئش بلەن ئوتكەرە. ئانك ئۇيىنە
بارسالك، يە ئوقب ئوترعانن، يە ئيسە ئىشك ئالدىندا برەر ئش ئشلەب
يۇرگەنن كورەسك.

Fig. 5. Tatar language in [8]. It includes initial, medial, final, and isolated forms.



Fig. 6. An early Soviet Tatar poster by A. Platunova. [9]



Fig. 7. An early Soviet Tatar poster by G. Arslanov. [9]

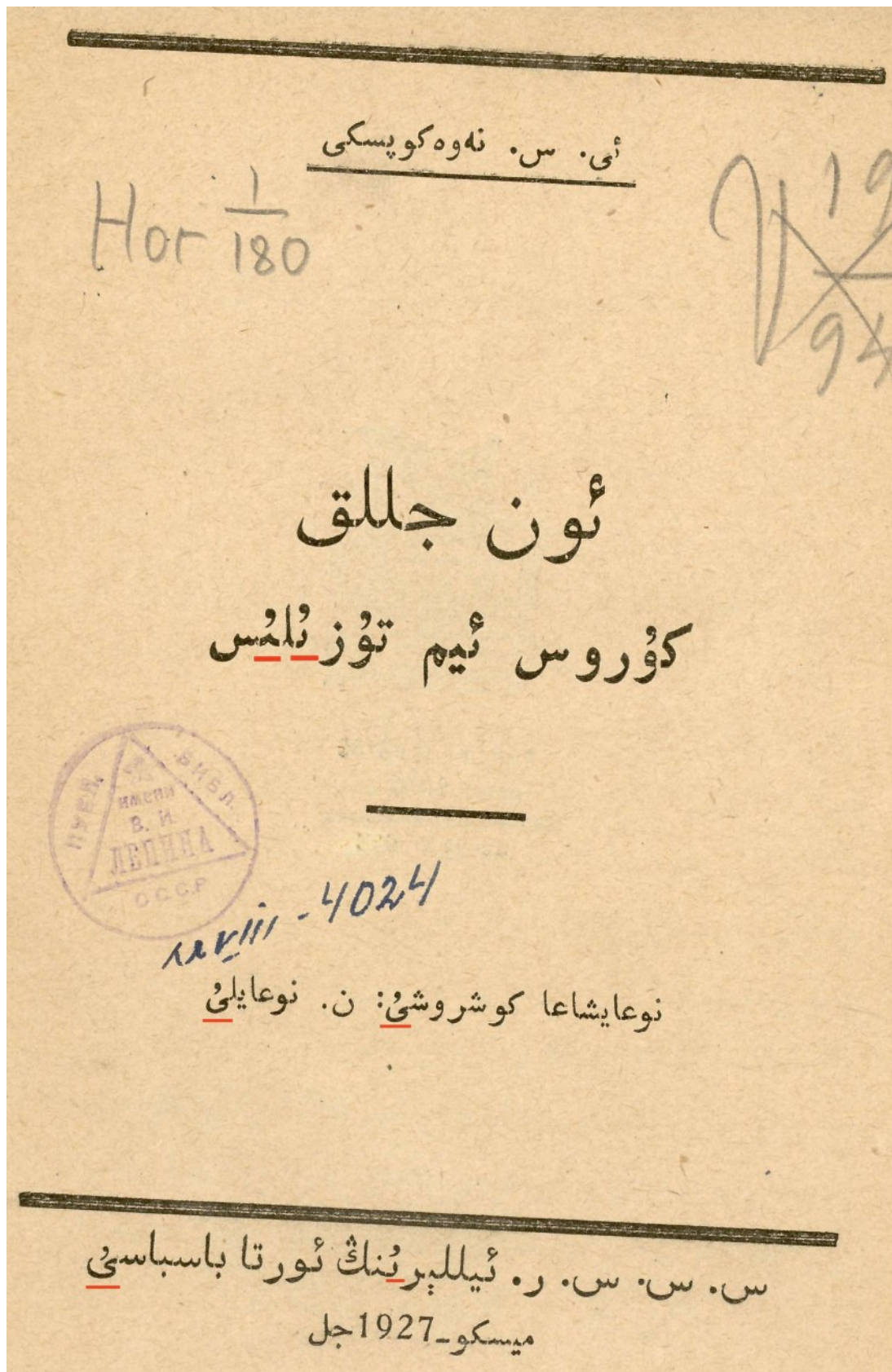


Fig. 11. A title page of a Nogai book. [18]

سوز باسى

بو ۋاقتقا دىيىم شىعارعان كىتەبلەر ئىمىز توراسىندا
ئوقۇشۇشلار ئىمىزدان تۇرلى قاتلار آلدق. بو قاتلاردى
جازۇشۇشلار، ئوز تىلمىزدى كىتەب شىعوغا شادلانې، بىزدى
بو ئىمىزدى قۇتلاپ جازادى، تىك، كىتەبلەردى جازودا
قۇللانغان ئىملەمىزدى تىز شىعارا آلماي آورسەنلىقلار
بىلدىرىدى. آشتارقانلى نوغاي ئوقۇشۇشلار ئىمىز: بىز بو
ۋاقتقا دىيىم تاتار ئەدەبىيەتى مەنىن تەربىيىلىنىپ كىلىدك،
قەزىردىدى ئەدەبىيەت، ماتېۋىيات ئوقۇشۇشلار ئىمىز تاتار
ئەدەبىيەت مەنىن پايدالانادى، سوغاي كورى، تىلمىزدى زاغىنا
كىلىسى، كىتەبلەر ئىگىزدى تاتار ئىملەسى مەنىن جازىب
شىعارساغىز ئىكەن، دىيىدى، تىك، تاتار ئىملەسىنىڭ (ئا)
ئورنى نا (آ) دى قوللانۇمۇمكىن بولماسىبا، ئولاي

Fig. 12. Nogai language in [19].

بۇرۇننان بىرلىكى كېلى ياتقان رىۋەلوتسىيى بايرامى
 بىرئىشى ماي بۇتن جىر جۇزىندىكى ئىمىيىكىشى قالقۇن ئورتاق
 بايرامى بولدى. بىرئىشى ماي بايرامى بارلق رىۋەلوتسىيى بايراملارنى
 آراسىدا ئۆزىنى آيرىم ئورن تۇتادى، ئىمىيىكىشى قالقۇن بۇ بايرامدى
 ئۆز آلدۇنا كۇتب آلدى. بو بايرام رىۋەلوتسىيى بايراملارنىڭ بىرى
 بولسادا، بو بايرامدى جاڭغى بايرام دېپ آتاب بولمايدى. كۆپ
 تۇرلى مېملىكىتىلەردى، بۇردىن سەيىتلىرى ئېلىندى، بىرئىشى
 ماي كۇنى بايرام ئىتلىمى باسلاۋعا 36 جىل بولدى. تىك
 بو 36 جىلدا بولغان 36 بايرام بارىنى بىردى تىگىس ئۆتكىزىلگىن
 جوق. بىزلىر، بىرئىشى ماي بايرامى، تىك توعۇز جىلدىن بىرلى
 جاسىرنامى، ئىمىيىكىشى قالقۇن تويى ئىتت ئۆتكىزىپ كېلى ياتىمۇ.
 رىۋەلوتسىيى كې دىم، بىرئىشى ماي بايرامى بىزلىر جاسىرن، قاشب
 پۇسب ئۆتكىزىدىگىن ئىدى.

Fig. 13. Nogai language in [20].

نىندى جۇعوسلىنى آورولار بولدى،
 ئۇلاردى قەيتىب بىلىنلىرى
 مۇنا مىسالان آيتايىق، قاي بىر كىشىلىرىدىن كوزلىرى ساودا بولسادا
 ئۇواق جازودى تانى ئالمايدى، ئول كىشىلىرى دوكتىرگى بارب كوزلارنىڭ
 آلدى. كوزلارنىڭ ئۇواق جازودى ئرى ئىتت كورسيتىپ باياعى
 كىشىلىرى بو جازودى ئوقى ئالدى.

Fig. 14. Nogai language in [21].

قاتن - قزلار ئاراسىدا ئىسلەۋدىن كەرەكلىكى.
 ئۆلگىن ئۆكتەبىر رەۋەۋەتسەيەسى رەسەيدەگى قالقىن
 تۇرمىسىن جاڭگارتىدى. باتشا قۇكومەنى تاراپىنان ئەزلىگەن
 مەللەتلەر (تاتار، قازاق، نوغاي ئەم باسقۇچلار) ئەركىنلىككە
 شەكىللەندى. مۇلىيەنلاغان ئىشى - ئەگىنىلەر سەۋەت قۇكومەتنى
 معاياقلاۋ ئىشە قاتتىق، جاڭگى سەسەيەلەزم، تۇرمىسىن
 قۇرب جاتىرلار.

Fig. 15. Nogai language in [22].

ئەرو جىرلار

نوغاي ئەرو (سەنەت) جىرلارى؛ قوس ئاۋاز، (1)
 شىك (2) (يار - يار)، مەيت (بەيت)، ئىنار دەگەن تۈرلەرگە
 بولۇپ بولىدۇ.

قوس ئاۋاز

I

قارايدار (3) مىنان قىزى كۆل

قوس ئاۋاز، ئاشتارقان نوغايلىرى ئاراسىدا ئەڭ
 قەدىرلى جىرلاردان سانالادى. ئول، قاپقاز نوغايلىرى
 ئاراسىدا سۆيىلىپ جىرلانادى.

Fig. 16. Nogai language in [23].

《塔塔尔文字母表》的字母与国际音标对照表

顺序	单写式	首连式	中连式	尾连式	国际音标
1	ا // آ	آ	ا	آ // ل	a
2	ه // ڤ	ڤ	ه	//	ε
3	ب	ب	ب	ب	b
4	پ	پ	پ	پ	p
5	ت	ت	ت	ت	t
6	ج	ج	ج	ج	dʒ
7	چ	چ	چ	چ	dʃ
8	خ	خ	خ	خ	x
9	د	د	د	د	d
10	ر	ر	ر	ر	r
11	ز	ز	ز	ز	z
12	ژ	ژ	ژ	ژ	ʒ
13	س	س	س	س	s
14	ش	ش	ش	ش	ʃ
15	غ	غ	غ	غ	ɣ
16	ف	ف	ف	ف	f

Fig. 17. Tatar alphabet in China in [24, p. 44]. (a) This is the first half of the whole table and it does not include ڭ.

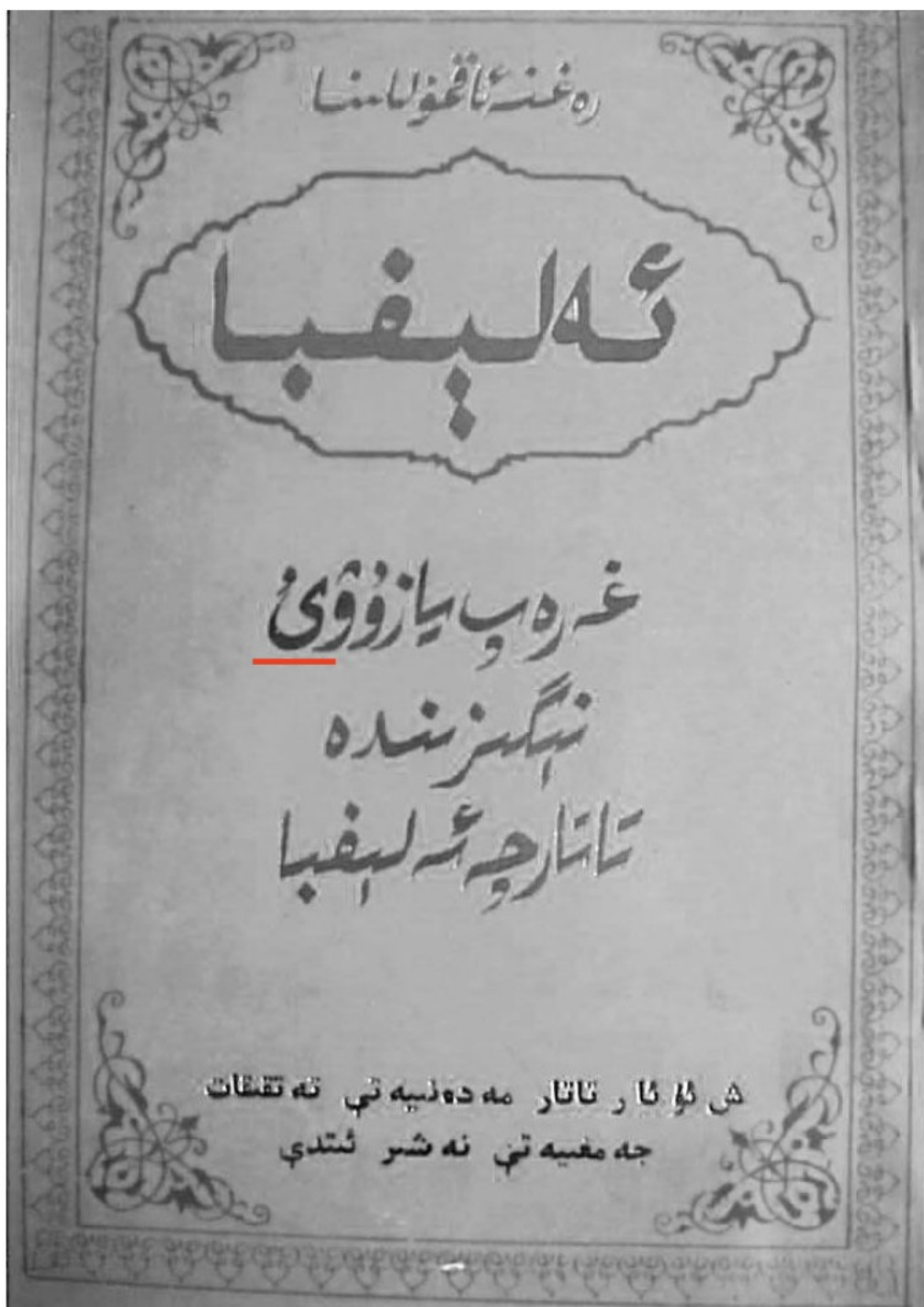
续表

顺序	单写式	首连式	中连式	尾连式	国际音标
17	ق	ق	ق	ق	q
18	ك	ك	ك	ك	k
19	گ	گ	گ	گ	g
20	ڭ	ڭ	ڭ	ڭ	ŋ
21	ل	ل	ل	ل	l
22	م	م	م	م	m
23	ن	ن	ن	ن	n
24	ه	ه	ه	ه	h
25	ئو	ئو	و	و	o
26	ئۇ	ئۇ	ۇ	ۇ	u
27	ئۆ	ئۆ	ۆ	ۆ	ø
28	ئۈ	ئۈ	ۈ	ۈ	y
29	ۋ	ۋ	ۋ	ۋ	w
30	ئې	ئې	ې	ې	e
31	ئى	ئى	ى	ى	i
32	ئى	ئى	ې	ى	ɨ
33	ي	ي	ي	ي	j

注:1.另有二个“半元音”,分别用“ﻻ”“ﻻ”字母表示。

2.本“字母表”于1941年,由新疆伊犁地区塔塔尔文化促进会制定,经伊犁地区教育局批准正式使用。

Fig. 17. (Continued.) Tatar alphabet in China in [24, p. 45]. (b) This is the second half. Note 2 in this figure states: “This ‘alphabet chart’ was formulated in 1941 by the Tatar Cultural Promotion Association of the Ili region, Xinjiang, and was officially approved for use by the Ili Regional Education Bureau.” The middle form may be a mistake.



塔塔尔族文识字课本（热娜·阿克莫、莉纳编，新疆维吾尔自治区塔塔尔族文化研究会出版）

Fig. 18. Tatar textbook used in China. The title is “Ələfba: Ğərəp yazuvı negizində tatarçə ələfba” (Ələfba: Ğərəp yazuvı negizində tatarçə ələfba)”. Reproduced from an image in [24, p. 86].

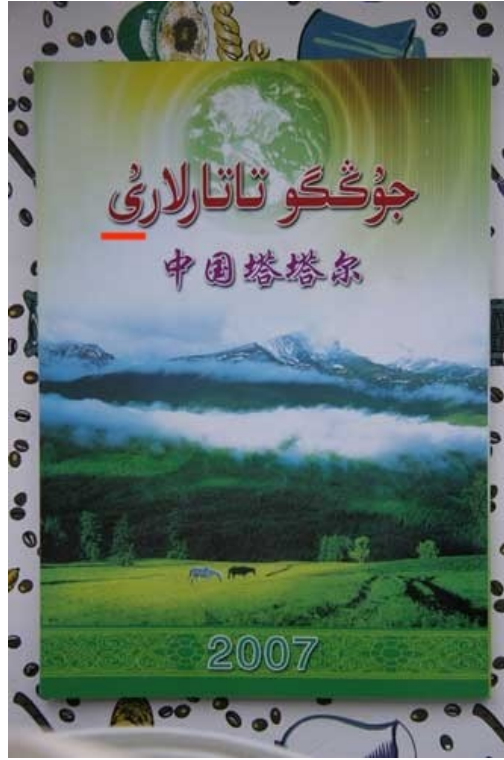


Fig. 19. A front cover of a book about Tatars in China (中国塔塔尔). The Tatar title is “جوڭگو تاتارلاری” (Cuñgo Tatarları). [25]

قاراغالفاق تىلىندەگى

ھەرف بېلىگى لەرى

ا ب ت ج چ د ر ز س غ ف
ق ك گ ل م ن و و و
ئ ي ھ ھ

بەر جۇغۇ جىكرما بەس ھەرف.

سۆزىلاتوغون داۋىسلا، ا و ئ ي ھەا غ
ق، ھەرفلەرى بەر قوللا جوۋان وقولادى، بول
ھەرف لەردىڭ بىردەن بىرەۋى ارالاسقان
سۆز بۇتۇنلەي جوۋان ايتىلادى.

ه گ ك ھەرفلەرى بەر قوللا جىڭچىكە
بولۇپ وقولادى. وننان باسقا ھەرف لەر
سۆزلەردىڭ اسلىنا قاراي قاي بىر جەردە
جوۋان، قاي بىر سۆزلەردە ەلجۇاز نەزىك
بولۇپ وقولادى.

Fig. 20. Karakalpak alphabet chart in [26]. The same chart appears in [1, Fig. 2].

چۇقاي بى دىڭ سۆزى

قان الدىنا بارغاندا
قان كەسبەدى تىلمىدى
قالق الدىنا بارغاندا
قالق تاييادى مىنىدى
ۈز ۈيۈمە كەلگەندە
ئەيتگە بەرسەن كۈنۈمدى

گۈدى ساقلاغان

بىر قاراغالىق دىڭ مايدانغا جىناپ قويغان فچەن كۈدى سىنىنەن
كۈندە بىس باۋ ون باۋ فچەنى ۈرلانا بەرگەن بىكەن وننان سوڭ
بىر تۇنى اڭلىب قۇرب تورۇب ۇرىنى ۇسلاپ الاين دىب ۇيۇق
لاماي گۈدىنى ساقلاپ وتىرغان بىدى.
مەيران قالغانداى ول تۇنى ۇرى كەلمەدى. بىلىمەن مۇنىڭ
اڭلىب تۇرغانىن بىلب تاپسالادىما ياماسا چىنى بىلەن دۇس بولمادىما
ئەدە بولسا ول گۈن ۇرى كەلمەدى.
تاڭ راۋان اتقاننان سوڭ گۈدىنىڭ بىغدا بۇيۇغۇب وتىرەپ
ۇيۇقلاپ قالدى.
مايداندا بارلاپ. اڭلىب تۇرغاندا ياماسا باسقا جەردەن-تالابى
كەلسەبى قايتىپ باراتقان بىر ۇرما دۇس كەلىپ مۇنىڭ قاتىدى
ۇيۇقلاغانىن كۆرەپ اقرىن غانا يىتىگەن سۇۋىرەپ الدىدا ۈتەگىتىدى.

Fig. 21. A page of a Karakalpak book [26]. The proposed character is not found.

بیرلەستىڭىز فۇنىكۇل دۇنيا مېتەنچىلارى
 Кара-Кант $\frac{4}{958}$ 188
 بىرىنچى بىلىم جولى 595

قارا خالفاق تىلىندە
 الېبېه

xxvi-8/83



جىناۋچى س. مەجىدوف.

•••••

بۇل بىر جۇل باسچى بالالار
 تارتار سىزدى دانالىق غا
 ادۋانى اباد قىلغانداي
 جۇل قالدىرماس قارالىق غا
 گىلىڭ لەر بايقاب قارايىق
 بىلگەن بىل بىلەن بىر بارايىق
 بىزدە بىلگەن بىل جولايق
 ماز بولمايق بالالىق غا
 بىلىم بولسا قولدا بارلىق
 ۋچباس جارلىق كىتبەس بايلىق
 ورنار بىرلىك تۇسبەس تارلىق
 قوچلا سارمىز الالىق غا

С. Мажитов.
 Букварь
 на каракалпакском
 языке.

56

Fig. 22. A page of a Karakalpak book [27]. The proposed character is not found.

مەكتەپكە

(بالالارغا)

جاڭا جاۋغان قارباركۇن كۆزى اچىق
 دۇنيا جارقىلداغان نۇردىڭ ءچىندە
 بىتلىرى دومبغوب تۇسلەرى قاچىق
 اداملار تور ءزغار سوردىڭ ءچىندە
 كۇن چىقىدى ۋاق بولدى مەكتەپ بارماڭا
 جاترسىز ۇيۇقلاپ جاقىن قالماڭا
 بىرتە بارتىپ تىزراق بىلىم الماڭا
 اسىق ازاڭشى اياز كۇنىڭ ءچىندە

كۈلكى

اي بىرنازار سەن نىگە وبكەلەيدىڭ اۋزىڭدى نەگە تومغايتا سالڭ?
 تۇنلە وتقا تۇسۇپ چاپانم كۇيۇپ قالىدى.
 نەگە جاماب المايسالڭ?
 ءينەم ساباغىم جۇق.
 كۆپ كۇيۇبىيە تىسىڭى ۇلكەنبە?
 جاڭاسى باراۋ تەڭ بىتەڭى بىلىن جەڭى جۇق.

Fig. 23. A page of a Karakalpak book [28]. The proposed character is not found.

Acknowledgment

The author is grateful to Naza Batyrovna Urakchieva, Head of the Information and Bibliographic Department at the State National Library of the Karachay-Cherkess Republic named after Kh. B. Bayramukova (Cherkessk), for kindly providing valuable information on Nogai Arabic-script publications of the Soviet period.

The author is grateful to the Russian State Library (Moscow), whose Reference and Bibliographic Sector identified its holdings of Nogai books published in 1926–1928, and whose Department of Interlibrary Communications and Remote Services kindly provided scans of them, which are used in this proposal.

The author is grateful to Bekmurza Gúlimbetovich Dilanov (Candidate of Philological Sciences, Associate Professor, Karakalpak State University), whose scientific-methodological research focuses on Karakalpak linguistics, for kindly providing scans of pages from S. Májitov's three Karakalpak-language Arabic-script textbooks.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	<i>Proposal to encode ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE</i>		
2. Requester's name:	<i>Takehiko Kyono</i>		
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>		
4. Submission date:	<i>2026-10-17</i>		
5. Requester's reference (if applicable):			
6. Choose one of the following:			
This is a complete proposal:	<i>✓</i>		
(or) More information will be provided later:			

B. Technical – General

1. Choose one of the following:			
a. This proposal is for a new script (set of characters):			
Proposed name of script:			
b. The proposal is for addition of character(s) to an existing block:	<i>✓</i>		
Name of the existing block:	<i>Arabic Extended-B</i>		
2. Number of characters in proposal:	<i>1</i>		
3. Proposed category (select one from below - see section 2.2 of P&P document):			
A-Contemporary <i>✓</i>	B.1-Specialized (small collection)	B.2-Specialized (large collection)	
C-Major extinct	D-Attested extinct	E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic	G-Obscure or questionable usage symbols		
4. Is a repertoire including character names provided?	<i>Yes</i>		
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>Yes</i>		
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>		
5. Fonts related:			
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Takehiko Kyono</i>		
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>Takehiko Kyono</i>		
6. References:			
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>		
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>		
7. Special encoding issues:			
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<i>No</i>		

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	Yes
If YES explain	<i>The character was previously proposed in L2/11-138 under the name ARABIC LETTER FARSI YEH WITH DAMMA ABOVE. It was not encoded because L2/11-201 considered it representable by the sequence <U+0649 ARABIC LETTER ALEF MAKSURA, U+064F ARABIC DAMMA>.</i>
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	No
If YES, with whom?	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	Yes
Reference:	<i>See proposal</i>
4. The context of use for the proposed characters (type of use; common or rare)	Rare
Reference:	<i>See proposal</i>
5. Are the proposed characters in current use by the user community?	Yes
If YES, where? Reference:	<i>Tatars in China</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	No
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	No
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?	No
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	No
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	No
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	No
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	

UnicodeData.txt

1 0892;ARABIC LETTER ALEF MAKSURA WITH COMMA ABOVE;Lo;0;AL;;;;;N;;;;;

ArabicShaping.txt

1 0892; DOTLESS YEH WITH COMMA ABOVE; D; YEH