Proposal to add Parkari letters to Arabic block

Date: November 1, 2001
Author: Jonathan Kew, SIL International
Address: Horsleys Green
High Wycombe
Bucks HP14 3XL
England
Tel: +44 1494 682306
Email: jonathan_kew@sil.org

A. Administrative

1.	Title	Proposal to add Parkari letters to Arabic block
2.	Requester's name	SIL International (contacts: Peter Constable, Jonathan Kew)
3.	Requester type	Expert contribution
4.	Submission date	November 1, 2001
5.	Requester's reference	
6a.	Completion	This is a complete proposal
6b.	More information to be provided?	No

B. Technical — General

1a.	New script? Name?	No
1b.	Addition of characters to existing block? Name?	Yes — Arabic.
2.	Number of characters in proposal	3
3.	Proposed category	A
4.	Proposed level of implementation and rationale	1
5a.	Character names included in proposal?	Yes
5b.	Character names in accordance with guidelines?	Yes
5c.	Character shapes reviewable?	Yes
6a.	Who will provide computerized font?	Jonathan Kew, SIL International
6b.	Font currently available?	Yes
6c.	Font format?	TrueType
7a.	Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	Yes
7b.	Are published examples (such as samples from newspapers, magazines, or other sources) of use of proposed characters attached?	Yes
8.	Does the proposal address other aspects of character data processing?	No

C. Technical — Justification

1.	Has this proposal for addition of character(s) been submitted before?	No
2a.	Has contact been made to members of the user community?	Yes
2b.	With whom?	Dr. Richard Hoyle (SIL linguist) has had extensive contact with the Parkari Language Committee, Parkari Community Development Programme, and others of the Parkari community
3.	Information on the user community for the proposed characters is included?	Yes
4.	The context of use for the proposed characters	Cultural, educational, medical, agricultural, religious, and children's books and other literature
5.	Are the proposed characters in current use by the user community?	Yes
6a.	Must the proposed characters be entirely in the BMP?	Yes
6b.	Rationale?	Contemporary characters in common use
7.	Should the proposed characters be kept together in a contiguous range?	No
8a.	Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	No
8b.	Rationale for inclusion?	N/A
9a.	Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?	Possibly, depending on calligraphic style, but characters are still distinct
9b.	Rationale for inclusion?	See below
10.	Does the proposal include use of combining characters and/or use of composite sequences?	No
11.	Does the proposal contain characters with any special properties?	No

D. SC2/WG2 Administrative

To be completed by SC2/WG2

1.	Relevant SC2/WG2 document numbers	
2.	Status (list of meeting number and corresponding action or disposition)	
3.	Additional contact to user communities, liaison organizations, etc.	
4.	Assigned category and assigned priority/time frame	
Other comments		

Proposal

History of the Parkari orthography

The Parkari language is spoken by about 500,000 people in the province of Sindh, Pakistan. Parkari has had a standardized orthography used in all written and printed materials since 1985. The orthography was standardized in 1983–4 by the Parkari Language Committee, made up entirely of members of the Parkari community.

Basis for Parkari orthographical decisions

The orthography is based on the Sindhi script, itself a modification of Arabic. Due to the difference in phonology between Parkari and Sindhi, three new letters were created to represent phonemes that are not found in Sindhi or Urdu. All three letters consist of an Arabic base form, already existing in Sindhi, with an inverted 'v' shape or circumflex-like symbol above the letter. They are as follows:

Phoneme	Description of letter shape	Example	Contrasts with phonemes	Written as
implosive dental /d/	dal with inverted 'v' above	ڐ	implosive retroflexed /d/	ڐ
retroflexed /l/	reh with inverted 'v' above	ڒ	normal /l/, retroflexed /r/	ڒ ل
voiceless /h/	heh with inverted 'v' above	ھ	voiced /h/	ھ

Two other Parkari phonemes (short /e/ and aspirated /z/) not present in Sindhi also need to be represented in the orthography. However, these are written using digraphs that can already be encoded as character sequences in the UCS (ئہ and زھ respectively); they do not require the addition of new characters.

Arabic has been regularly adapted in the past to match the phonology of other languages, e.g., for Farsi, Urdu and Sindhi. Digraphs have sometimes been used for aspirates (e.g., in Urdu). Other adaptations have involved modifying existing letters using superscripts or subscripts, typically using different dot combinations with a single base letter, or using the same superscript symbol on different base letters. Due to the frequent use of dots in Sindhi, the Parkari Language Committee chose to modify letters by adding a single superscript symbol to different base letters already present in Arabic (comparable to the use of a small TAH mark to create the Urdu retroflex letters).

Use of the Parkari orthography

The Parkari orthography has been used in published materials since 1985 and is accepted and used by all members of the community, as well as by outsiders working in the Parkari language. The orthography is used in printed books, in a community magazine, in formal education in rural schools, in cultural materials, on cassette covers, on cards, and for personal letters. As well as other organizations and individuals using this standardized orthography, it is used by the Parkari Community Development Programme (formerly Parkari Audio Visual), a Parkari organization founded in 1996, which produces and publishes Parkari books and promotes Parkari literacy.

Examination of the examples will show that there is some variation in the shape of the inverted 'v' above the Parkari letters; some writers and fonts use an angular form, similar to a circumflex accent, while others use a more rounded form, similar to a breve accent. This is purely a stylistic difference, not a meaningful distinction.

Amount of Parkari literature

There are over 100 books published in Parkari, ranging in size from children's Early Reading Books with only a few pages, to educational and medical books with some 30-50 pages, and including cultural and religious books with several hundred pages. These books cover a wide range of topics including cultural, educational, medical, agricultural, and religious materials, as well as children's stories. There is also a regular community magazine. Books have been published by the Parkari Language Committee, Parkari Audio Visual / Parkari Community Development Programme, the Catholic Diocese of Hyderabad, and Pakistan Bible Society.

Characters proposed to be added to Unicode

Three additional characters are required in the Arabic block of Unicode. They each share the same properties and joining behavior as the basic Arabic letters on which they are based. It is unimportant exactly what Unicode values are allocated for these letters, though they should be kept with the rest of the Arabic script. One option would be to use the three remaining code points in the U+06Ex..06Fx columns; another would be to allocate codes in the U+065x column.

Glyph	Suggested USV	Character name	Join type	Join group
ꗻ	U+06EE or 065D	ARABIC LETTER DAL WITH INVERTED V	R	DAL
ꗼ	U+06EF or 065E	ARABIC LETTER REH WITH INVERTED V	R	REH
ꗽ	U+06FF or 065F	ARABIC LETTER HEH WITH INVERTED V	D	KNOTTED HEH

In some hand-written forms of Arabic script, a combination of three dots pointing upward (as seen on U+0698, for example) may appear as a single stroke in the form of an inverted 'u' or 'v', similar to the mark added for the Parkari letters. This might appear to suggest that the Parkari DAL WITH INVERTED V and REH WITH INVERTED V should be unified with U+068E and 0698 respectively. However, this would be incorrect. Although the Parkari orthography itself does not require either U+0698 or 068E, there is a strong possibility of U+0698 occurring in Farsi and Urdu loanwords. The form with three dots, representing Farsi JEH, would then contrast with the Parkari letter with the inverted 'v', representing the retroflexed /l/. Similarly, if U+068E occurs in Sindhi text, this should contrast with the Parkari letter.

It is also significant to note that the use of the inverted 'v' mark was a deliberate choice on the part of the Parkari Language Committee, taken in view of the fact that multiple-dot combinations are already so frequent in the Sindhi script they took as a basis. Rather than use yet more letters with multiple dots, such as by 'borrowing' the JEH letter from Farsi/Urdu, they preferred to choose a visibly distinct form, and this choice has become the accepted standard throughout the community. In all the published examples of Parkari literature, it can be clearly seen that patterns of distinct dots are written or printed for all the standard Sindhi letters, while the three Parkari additions have their unique mark instead. While it may be true that in some handwriting, a three-dot pattern might 'degenerate' towards the form of the Parkari mark, it is not appropriate to unify the distinctive Parkari letters with three-dot versions.

Samples of published Parkari literature

The following sample pages are scanned from Parkari books published in Pakistan during the period 1985–2001. In each case, a few examples of the three special Parkari letters are circled for ease of reference.

سنڌي	پارکري	سنڌي	پارکري	سنڌي	پارکري
ا	ا	ب	ب	ب	ب
پ	پ	ت	ت	ت	ت
ت	ت	ٿ	ٿ	ث	ث
پ	پ	ٯ	ٯ	ج	ج
چ	چ	ڄ	ڄ	ح	ح
خ	خ	ڇ	ڇ	د	د
د	د	ڈ̂	—	ڊ	ڊ
ي	ي	ڌ	ڌ	ر	ر
ر	—	ڙ	ڙ	ز	ز

سنڌي	پارکري	سنڌي	پارکري	سنڌي	پارکري
ش	ش	س	س	—	زھ
ط	ط	ض	ض	ص	ص
ڠ	ڠ	ع	ع	ظ	ظ
ک	ک	ق	ق	ف	ف
ڳ	ڳ	گ	گ	ک	ک
ل	ل	ڱ	ڱ	گھ	گھ
ٽ	ٽ	ن	ن	م	م
—	ۿ	ه	ه	و	و
		ي	ي	ء	ء

٣٥   ٣٤

Transfer Primer, published by Parkari Language Committee (1985) 36 pages (pp34–35 shown)

These charts show a comparison of the Parkari and Sindhi alphabets. Note the three Parkari letters with inverted 'v' above, which do not have Sindhi equivalents. (`The /zh/ digraph also lacks a Sindhi equivalent, but this combination does not need to be encoded as a distinct character in the UCS—it can already be adequately encoded as a character sequence.)

ويل .

:ا من موئين ڋيپ جلاوو، پريئو جي مارا
:ا من موئين ڋيپ جلاوو :ا

1 . :ا ڱهور انڌارو مارڳ نا ڋيکون
ست نيڱري نئہ کم کرينَ اوڙ کون :ا
:ا هاٿ زهالي ليو، او مارا سوومي :ا
پريم رو پنٿ ڋيکاڙو پريئو جي مارا
من موئين ڋيپ جلاوو .

ويل .

2 . :ا آوينَ رو مارا جيوَن موئين
پلٿو مارون پيئُون نيٿون موئين :ا
:ا هُون سون تمارو آپرو پريئو جي :ا
هِيئي هٿيو ھماوو، پريئو جي مارا
من موئين ڋيپ جلاوو .

ويل .

3 . :ا پروڻ ڋيڌو منکون ھارُو
کروسي سڙيا پايِيون ھارُو :ا
:ا لوئي رئي ڏار جي ووڋئي پريئو جي :ا
اوَئا موئين امون نئہ جهيلاڙو،
پريئو جي مارا
من موئين ڋيپ جلاوو .

ويل .

Hymnbook, published by the Catholic Diocese of Hyderabad (1985) 734 pages (p 352 shown)

پوهلئہ ان کلکوڑي هاٹ ہونڈین کِیڈو
"او پَوِیتر سنت، امارو ٹینصلو کرو."
مِینڑئہ کِھڈو
"مارون ڈیکروں، هوں هَوئہ ڈوکرو ٹي گِھوه.
هوں نئہ گِھٹ هومِیڑیامَ آوئہ.
تمي ہي آوین
مارا کون مۄئین واتون همِیڑاوو."

Jungle Stories, published by Parkari Language Committee (1988) 48 pages (p 39 shown)

۱٦۵

کُونٹا رئي لاکڑئي لٹ پٹ ٹئه،

پوِٹي رون سونٹون هُون نر ان نارئي زُوئا ٹئه.

(کوڈ: مِنوز، زڑوز)

۱٦٦

کونٹون رئي اُٹل پاٹل، پونڑوں رو ڈکاڑ،

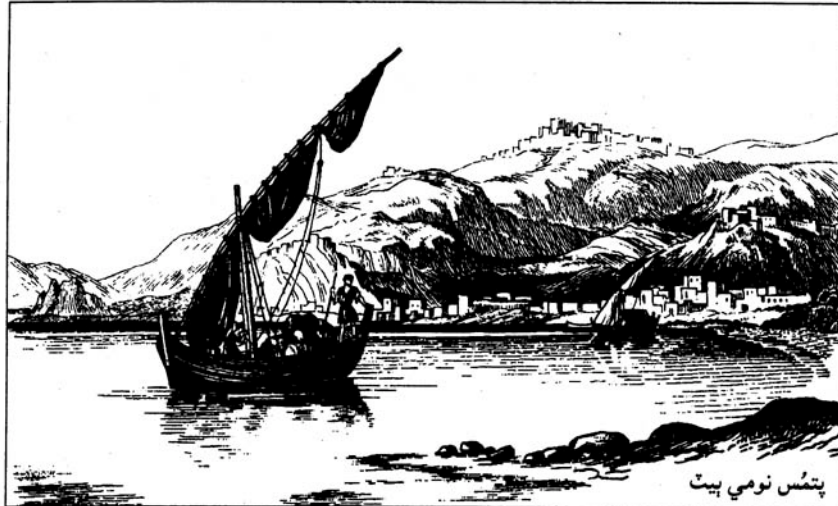ٹڑیا موٹهئہ ٹڑ را، هوپاري کاٹا واڑا.

(کوڈ: اڈیوز)

✿

۱٦۷

کوري موري پاپڑیا،

اُسکاوَي زو راٹئي سوڑ.

(کوڈ: مِنوز)

✿

۱۵۲

Proverbs and Riddles, published by the Catholic Diocese of Hyderabad (1990) 185 pages (p 152 shown)

اِیسُو رو یُوحنا نئہ ڈیکاڑل

ڈرشن

پتمُس نومِي بيٽ

① اِیئا کِتاب مِ اُو واتون لکل سئہ جي اِیسُو کرِیست ڈرشن ڈِینَ ساوي ڈیکاڑِیھ. هائي واتون اِیسُو نئہ پِچُھونَ را پاها ٿِني زڑي زم اُو آپرون شیواڈاریون نئہ ساوي ڈیکاڑِئہ، جي کِي کِي واتون ضرُور جلڈِي مِ ٿیشي. هائي واتون اِیسُو آپرا آکاشِي ڈُوت نئہ میلینَ مُون آپرا شیواڈارِي یُوحنا نئہ نِیشونون ٿي ساوي ڈیکاڑِي. ② ان پِچُھونَ را وچن ان اِیسُو کرِیست را پارام زتي واتون مٿِین ڈیکیھ، اُوئُون ہِڈیون واتون را پارام هُون هاک پِرونھ. ③ پاگُواڑو اُو سئہ جي پِچُھونَ را اِیئا نیاپا ري واتون پٽِینَ همیڑاوئہ، ان پاگُواڑا اُو سئہ جي اي واتون هُومیڑئہ، ان اِیئا نیاپا ریون لکل واتون مانہ وئہ. کمِ تو اِیسُو رئہ پاسا آیا رو ٹوٹو ڈُوکٹرو سئہ.

The Parkari New Testament, published by Pakistan Bible Society (1996) 119 pages (p 1030 shown)

هفتو 4   مضمُونَ: گُروم

انگارو: هيكا گُروم مِ بي لينهڑزيون را وڻ هتا. هيك ڈي گُروم رِي منكي لينهڑزيون را هيزا آن وڻ هوكيا. هيارون هازون، تمي همِيزاوو جي لينهڑزيون را كتا وڻ ٿيا؟

اريا: هيك موٽو گُروم هتو. هايا گُروم مِ هيك موٽئي اسپتال هتئي. هايا اسپتال مِ سو ڈاكٽر هتا. هي ڈاكٽر هيزا آيا. هيارون هازون، تمي همِيزاوو جي اسپتال مِ كتا ڈاكٽر ٿي گيا؟

كميس: هيكا گُروم مِ نَوَ نلكا لاگل هتا. اُرٽون موٽين ٿي هونس نلكون رو هوٽي كارو هتو. هيارون هازون، هوئہ تمي همِيزاوو جي ياتي كتون نلكون رو هوٽي ميٽو هتو؟

هفتو 5   مضمُونَ: هاڙل زناوَر

انگارو: هيكا گُروم مِ چار مِينزئي هتي. هيك ڈي هايا گُروم مِ هيك هيزئي مِينزئي آوٿي. هيارون هازون، تمي همِيزاوو جي كتي مِينزئي ٿيہ؟

اريا: هيكا منك را سو ٿاها هتا. هيك ڈي اُو شيكار مانہ گيو، ان هيزا ترڻ ٿاها رچ ہِ زهالي لِيڈا. هيارون هازون، تمي همِيزاوو جي كتا ٿاها ٿيا؟

كميس: هيكا گُروم مِ نَوَ كُوترا هتا. ان هايون نَوون موٽين ٿي چار كُوترا ميا هتا. هيارون هازون، هوئہ تمي همِيزاوو جي ياتي كتا كُوترا ٿازا ٿورا هتا؟

28

School Mathematics Book 1, published by Parkari Audio Visual (2000) 31 pages (p 28 shown)

"آئون لُوسِي تروٽ، انگلينڊ ۾ آڪسفورڊ جي رهڻ واري آهيان.
۽ جارج بُوٽ تروٽ جي عزيزن مان آڪري هڪ واري آهيان.
جان بُوٽ تروٽ 1830ء ۾ پيدا ٿيو، ۽ 1875ء ۾ وفات ڪري ويو.
1847ء ۾ 17 ورهين جي عمر ۾ برطانيه جو سپاهي ٿي هندوستان ويو.
هُن تمام جلدي ترقي ڪئي. ۽ جڏهن 1857ء ۾ انڊيا برطانيه جي خلاف جنگ ڪئي
تڏهن هو سنڌ ۾ سياسي سپرينٽندنٽ هو.
نگرپارڪر ۾ هن جي خلاف وڙهن وارن مان هڪ پارڪريو ڪولهي هو.
جنهن جو نالو رُوپلو گُوٽيل هو.
هو گهٽي بهادري سان وڙهيو ۽ مريو. تڏهن به هُو اڄا تائين پارڪري قوم ۾ هيرو آهي.

1999ء ۾ پُونم پاسڪل پرمار انگلينڊ آيو،
۽ هن مُون کي اها رُوپلي گُوٽيل جي سچي ڳالھ ڪري ٻڌائي ۽ فوٽو ڊيڪاريا جيڪا هن ڪتاب ۾ آهن.
هن صدي جي نئين شروعات ۾ هي موقعو آهي ته پراڻا وير ختم ڪريون.
آءُ پارڪري ڪولهي ماٿمن جي لاءِ ۽ سنڌ جي رهواسين ڇي لاءِ
نيڪ خواهش رکي امن ۽ صُلح جو نياپو ڏيان ٿي.

صحيح
لُوسِي سي. ايس تروٽ   (ايل. سي. ايس تروٽ)

"هُون لُوسِئي تروٽ، انگلينڊ ۾ آڪسفورڊ رئي ريا وازئي سون،
ان جارج بُوٽ تروٽ رون عزيزون موئين ٿي آڪري هيڪ سون.
جارج بُوٽ تروٽ 1830ء ۾ زلميو تو ان 1875ء ۾ مري گيو تو.
1847ء ۾ هترئہ ورهون ري عُمر ۾ پرطانيہ رو سپاهي ٿين اُو هنڊوستون گيو.
اُوئہ جوم جلدي ترقي ڪرئي ان زئيون 1957ء ۾ انڊيا پرطانيہ رئہ ڪلاڦ وڊيو
تئيون اُو سنڌ مانھ سياسي سُپرينٽنٽ هتو.
نگر ۾ اُوئا رئہ ڪلاڦ وڊيا وازون موئين ٿي•
هيڪ پارڪريو ڪوزي هتو جيا رو نوم رُوپلو گُوٽيل هتو.
اُو جوم بهاڌري ٿي وڊيو ان مروئو، توئي پڻ اُهزون هُوڊي پارڪري نات ۾ هيرو سئہ.

1999ء ۾ پُونم پاسڪل پرمار انگلينڊ آيا،
ان اُوئي مُون نئہ هائي رُوپلا گُوٽيل رئي ٻڌئي وات هميزاوئي ان قوٽو ڊيڪاڙي جيا ايٿا ڪتاب ۾ سئہ.

ايٿا صدّي ري نوي شروُعات ۾ اي موقعو سئہ جي زُونا وير رو انت ڪرون.
هُون پارڪري ڪوزي منڪون هارُو ان شنڌ رون ٻٽون ريواشيون هارُو
نيڪ آش راڪينَ شونتي ان هڪ رو نياپو ڏيونٿ.

صحيح
لُوسي سي. ايس. تروٽ   (ايل. سي. ايس تروٽ)

Cultural History of Parkar (in Sindhi), published by Parkari Audio Visual (2000) 251 pages (p 5 shown)

The lower half of this page contains a Parkari translation of the Sindhi text in the upper half.

ہار تائوکئہ ڈارئہ کم کرینَ شِیکئہ

ایوئہ نمُونہ شِیکاڑوو زم شِیکل واتون **همیش هُوتري** رئہ.

⊕ پیٹیا واژون مانھ کاس ڈیونَ:

پیٹیا رئی هیک کاس عمر هوئہ. پیٹیا واڑا نہ جی کئین پیٹاوا کروه ریا اُو هائی پیٹیا جیوو سئہ (عمر پرهوٹئہ) پیٹیا ہ ہار رو شونق سئہ؟ پیٹیا واژون رو مونَ راکو. پیٹیا واژون هارُو حفاظت هوئہ. هایا عمر ہ جیوو پن تجربو ٹیشی اُو همیش هُوتري ري زاشي. راجُي کُٹي جی کئین شِیکشي اُوٹا ٹي گهٹو کئین پیٹشي.

⊕ هوشیار:

ہار کوئی پن کوم کریا ٹي شِیکئہ. آزاڈئي ان تجرِپو پیٹیا ان شِیکیا رئہ پاهٹہ لي زائہ.

⊕ ڈورووٹ رو توڑ:

سڈئي ڈان کروو. ڈورووٹ رو توڑ گُوتوو. ہار نہ هم ٹائوکو لاگِشي ان اُو ٹائوکئہ ڈارئہ پیٹي هیکشي.

⊕ پیڑا تینَ کوم کروو:

ویچار ان کیال هُومیڑیا ٹي شِیکئہ. ان پیڑا تینَ ڈورووٹ رو توڑ گُوتوو ان هم پیڑا ٹائوکئی شِیکومٹ زڑئہ ان ٹائوکو ڈوٹ پڑئہ.

⊕ آقُوکو آپ کینک کروو:

ہاژون رو وِشواس پکر ٹئہ زئیون اُو کینک کریا ہ کامیاب ٹئہ ان کیٹل کوم زئیون ٹائوکئہ ڈارئہ کرئہ تو اُو وڈارئہ پروسا ٹي شِیکئہ.

۳

Teacher's Handbook, published by Parkari Audio Visual (2001) 138 pages (p 3 shown)