

John Benjamins Publishing Company



This is a contribution from *Written Language & Literacy* 15:1
© 2012. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

The development of structural characteristics of Brahmi script in derivative writing systems

Liudmila L. Fedorova

Russian State University for the Humanities Moscow, Russia

Indian writing systems reveal a great variety of graphic forms, proceeding from the unique source which is Brahmi script. These graphic forms render structural oppositions developed in phonographic writing systems which stem from Brahmi. This paper aims to highlight the changes scripts underwent to satisfy demands of language structure. The comparison raises the issue of the complexity of writing systems.

Keywords: writing system; grapheme; typology; Brahmi; abugida; alphasyllabary; akshara; diacritic; graphon; vowel differentiation

1. Introduction

The typology of phonographic writing systems can be based on different (but interrelated) grounds – structural or functional. This distinction can be compared to the terms *paradigmatic* vs. *syntagmatic* used by Gamkrelidze (1994) with regard to writing systems and, particularly, to alphabetic systems derived from Greek. These terms are quite appropriate for the discussion of the Indian material in the present analysis. And, in general, they correspond to the structural and functional perspectives of research emanating from Cartesian and Gerderian approaches, as Hymes (1974: 78–79) notes.

The structural criterion refers to the linguistic nature of the basic corresponding unit of a written sign (phoneme, syllable, morpheme, or word). The functional criterion refers to the syntagmatic use of these signs, which can correspond to a sound, a mora, a syllable, a word. As phonographic writing encodes spoken language, there can be different correspondences between structural and functional units.

The term *grapheme* will be used in this paper to designate the independent full-formed unit of linear scripts which refers to a linguistic unit. As full-formed signs written within the vertical space of lines, graphemes are opposed to diacritics

as these are generally placed above or under the lines. As independent signs, graphemes have their own linguistic referents and are thus opposed to sub-graphemes, namely to both diacritics (even when placed in line) and parts of ligatures. As signs with a linguistic referent (referring to linguistic units), they are opposed, firstly, to punctuation marks (signs of division and integration) and, secondly, to diacritics that refer to features or characteristics of linguistic units and form modifications of invariant signs. Parallel to any other linguistic unit, a grapheme can have its allographs. Finally, due to their structure, graphemes can be elementary or compound, formed of several meaningful graphic elements composing a unit.

I am inclined to interpret Brahmi as belonging to the syllabic type of phonographic writing, both in structure and in functioning. Although according to Daniels' typology, Brahmi belongs to abugida writing which is different from syllabic writing, other authors highlight that Brahmi has a syllabic nature: "All Brahmi-derived writing systems are syllabic, i.e. characters represent consonant-vowel syllables" (Wali et al. 2009: 162). Bright (1996: 384–391) refers to such systems as *alphasyllabaries*, and so does Sproat (2006: 45):

"The Brahmi-derived Indic scripts occupy a special place in the study of writing systems. They are *alphasyllabic* scripts [...], meaning that they are basically segmental in that almost all segments are represented in the script, yet the fundamental organizing principle of the script is the (orthographic) syllable."

Since there is no complete agreement on classification terms,¹ I propose a rather different, semiotic view on the hierarchy and distribution of the types of writing systems, taking into account the referent of a writing unit. (cf. for the whole typological scheme Fedorova 2011)

Consequently, based on the linguistic nature of graphemes, I discern four main functional types of syllabic systems:

1. Different elementary graphemes represent different syllables (CVC, (C)V, VC, CCVC and others) as inseparable units; such systems historically developed from logosyllabic (morphosyllabic) writing, or were recently standardized, like the Yi; this type can be labeled as the *Yi* type.²
2. Different elementary graphemes represent only simple moras ((C)V, (V)C): the *Kana* type (following the Japanese model).³
3. The single grapheme represents syllables with the same consonant, accompanied by different vowels or by their absence (C^x) as an inherent characteristic of a syllable: the *abjad* type (to use Daniels's term (1996: 3–17)). Structurally, it can be described as a consonant alphabet since only consonants have a visible expression. Functionally, it can be seen as syllabic writing since it allows graphemes to be read as syllables (according to Gelb).

4. The single grapheme stands for syllables with the same consonant, whereas its modifications differentiate vowels (C^v): the *abugida type* (Daniels' term). Structurally, this system can be regarded as an alphasyllabary (Bright's term).

Notionally, it is possible to discern a fifth model with the vowel grapheme being differentiated by a consonant (V^c). Pahawh Hmong script comes closest to it (Ivanov 2004: 25; Rogers 2005: 260–263): though usually there are two graphemes for a syllable CV in this writing system, their order in script is inverse – VC. This may show a subordinate role of the consonant in a syllable block in which the vowel also has a tone function; so Pahawh Hmong combines properties of the alphabetic and the syllabic type – it has full-sized letters for vowel and consonant phonemes, and its syllabic units (formed as non-linear) follow each other in a linear order.

Each type is based on its own principle, or mode, of writing. When speaking about principles, or modes, of writing I mean the common rules, which underlie the way to relate speech with graphic symbols. The second type is opposed to the first one in the way complex sound sequences are divided in parts; the third and the fourth type of scripts divide the sound sequences further into subordinate parts – a consonant as a head and a vowel as a dependent element, a specific or unspecific characteristic of a syllable – and represent them in writing as subordinate vowel marks in an abugida or as a zero-sign corresponding to an unspecific vowel or its absence in an abjad (the zero-sign can be substituted with *matres lectionis* or with diacritic marks of vowels, cf. the Tiberian system). Under this conception, an alphabet makes the relation between the consonant and vowel components coordinative. I do not attempt to trace the historical evolution of script but rather I am trying here to give a semiotic perspective of its improvement that reveals in adaptation of a writing system to a language (the most common way in the development of writing systems). In fact, such adaptation deeply depends on the structure of a language – its phonological, morphological, and lexical values, as Pandey (2003: 41–61) notes.

2. The structural oppositions of the Brahmi writing system

Brahmi script has the form of an abugida, with an akshara (an orthographic syllable, a *syllabogram*) being the main item. The akshara is a traditional name for an autonomic grapheme representing a C^v syllable which refers to Ca in its alphabetic form but can denote other vowels in the forms modified with diacritic marks.

The abugida systems are already the next stage of differentiation of the abjad – a stage of vowel differentiation (yet, the Japanese Kana, belonging to the second functional type, is not a precedent, but a deviated pattern; Coulmas (1996:252) remarks on the Indian influence on its structure).

Regarding its internal organization, the Indian script is a great innovation: it allows representing the phonological structure of a word quite distinctively. Derivative scripts maintain and develop the main structural oppositions:

1. between CV and V: in the development of modes of representation of independent vowels
2. in the differentiation of short and long vowels
3. in the introduction of the opposition C^v : C^o (presence of a vowel vs. absence of a vowel), which enlarges the paradigm of a syllabic sign
4. in the integrated representation of consonant clusters
5. in the introduction of a system of tone differentiation
6. in the aspect of text division – marking a word, a phrase, or other units.

These structural oppositions make use of some graphic devices, the main ones of which are diacritics, ligatures, traits of integration, and marks of division. These devices differentiate signs and improve the writing system as a whole. The paper aims to specify some ways and modes of differentiation. In general, I aim to reconstruct the akshara's grammar, which is interpreted in accordance with the following idea: "If a sign system has rules how to form complex signs and how to combine signs to form larger sentences, it has a grammar, besides the lexicon" (Neef 2010:231); with respect to a writing system, I take the set of its graphemes to be its 'lexicon'.

3. Paradigmatic systems and their functioning: Alphabetic and abjad principles of writing

Put into historical perspective, Semitic script was a great precedent (relative to Brahmi) invention – not only of a protoalphabetic system that had given birth to the Greek alphabet, but also of a mode, or principle of writing, recently named *abjad*. In order to distinguish between the paradigmatic system and its functioning, I prefer to use the term *alphabet* for a paradigmatic system of graphemes which:

- a. correspond to the sounds – phonemes or syllables of a given language,
- b. are ranged in a stable order,⁴
- c. usually have appropriate names – which makes them *letters*.⁵

In its functioning, such a system can follow an alphabetic principle or one of the syllabic principles – *Yi*, *kana*, *abjad* or *abugida*, as introduced above. According to the *alphabetic* principle, each phoneme should be expressed by a full-sized grapheme. The *abjad* principle, on the other hand, presupposes the use of graphemes in the syllabic meaning C^x (where C designates a specific consonant and x as a zero-sign an unspecific vowel – e.g. we follow the *abjad* principle in reading consonants in the alphabetic row: /bi/, /ci/, /di/, but /jey/, /key/ with different vowels, though they are not written.). The *abjad* mode of writing is largely defined by the phonological, morphological, and lexical structure of classical West Semitic languages, where:

- a. a vowel could not open a syllable, so it was regarded as a dependent characteristic of a syllable;
- b. a vowel was variable in word-formation, so it was not a constant characteristic of a root.⁶

Yet, a vowel accompanies a consonant in these languages, and, therefore, it is a regular and implied characteristic of a consonant or a syllable as a pronunciation unit as well.

4. The abugida principle of writing: Vowel differentiation

The abugida principle was probably discovered in writing of Indian Prakrits, in which the structure of a syllable, which could be CV or V, presupposed two demands:

- a. Since a vowel could correspond to a syllable, it needed a special designation.
- b. Different vowels needed different designations.

4.1 Kharosthi

A possible first step towards systematic vowel differentiation can be seen in Kharosthi⁷ (used mostly for Gāndhārī, a northwestern Prakrit dialect) which has graphemes for independent vowels. Still, all of them can be regarded as variants of an invariant grapheme for /a/, as if the invariant form represented a consonant, e.g. ⟨'⟩ for a glottal stop, with an inherent vowel /a/ (as in other invariant graphemes) or an empty-consonant, so the whole alphabet (*arapachana*) could have a uniform structure of graphemes – CV. The device that makes it function as an abugida is a minor graphic element for vowel differentiation – I shall call it a *graphon*. A graphon in Kharosthi is adjacent or tangent to the main sign within its

own space in the line, and it produces modifications of the invariant grapheme. It may change direction or the point of contiguity in accordance with the outline of the grapheme, yet it preserves its distinctiveness. Thus, the akshara-grapheme receives the internal structure by an invariant base and a variable graphon. The paradigmatic structure of such a system has two components: an alphabet of Ca graphemes⁸ (the invariant shapes) and a matrix of their modifications with graphons.

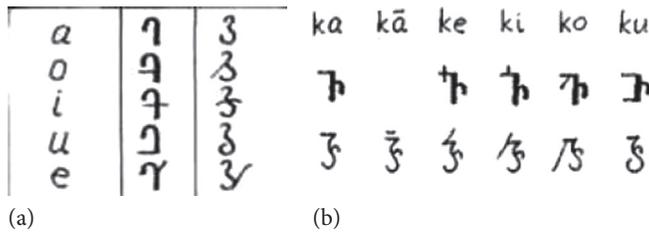


Figure 1. (a) Independent vowels in Kharosthi (two styles); (b) paradigm of vowel differentiation for a syllable Ka⁹

4.2 Brahmi

In Ashokan Brahmi, the differentiation of vowels goes further. Each vowel receives its own shape in an independent position, not on the base of a common invariant. It is a more developed abugida than Kharosthi; however, it is less uniform. It has two classes of graphemes: CV (a class for syllables) and V (a class for vowels). The latter can be conceived both as a class of syllables or of phonemes since the vowels have no modifications common with those of syllable graphemes. In addition, vowels can have their own modifications for the opposition of a short or a long sound. In this way, /a/, /i/, and /u/ obtain modified forms for long phonemes (in Kharosthi, only the long /a/ of a Cā syllable can be marked).¹⁰ Thus, the whole system corresponds to the phonological structure of Pracrits.

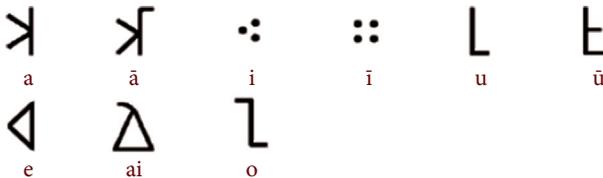


Figure 2a. Independent vowels in Brahmi

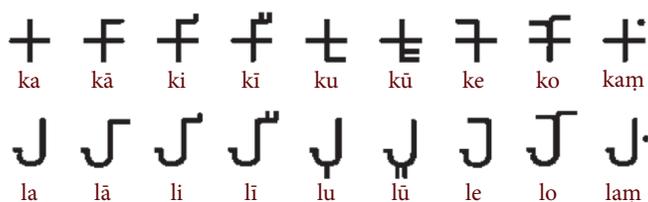


Figure 2b. Paradigm of vowel differentiation for Ka and La in Brahmi¹¹

Vowel differentiation of a syllable in an abugida is not confined to the choice of a graphon designating a concrete vowel, but it presupposes sequential operations of deleting /a/ from Ca and then adding a different vowel. Thus, a vowel mark, a graphon, is an operator that creates a new unit. The nature of the main grapheme would otherwise be uncertain: it may be understood as a sign of a syllable, an akshara, or as a sign of a consonant, a letter. In fact, in Ashokan script such uncertain uses can be found, but most graphemes in a text have a syllabic nature.

The role of graphons in the modification of vowels (V vs. V:) and of consonants (C^{v1} vs. C^{v2}) is similar. Thus, they can be conceived as the operators of (almost) the same order in contrast to the dots (*anusvara*), marking a secondary pronunciation feature.

Therefore, the paradigmatic structure of the Brahmi script has three main components: a system of vowel aksharas, a system of invariant Ca aksharas, and a matrix of their vowel modifications.

4.3 Tamil Brahmi

In early Tamil Brahmi, the direction of derivation differs: it tends towards the alphabetic writing. The invariant form of a grapheme corresponds to a pure consonant, while a syllable Ca is marked with a graphon (like the syllables Ci, Cu, etc.). If the basic form represents a pure, unvocalized consonant, it is not a syllable but a phoneme which forms the basis of the system (vowels can also be perceived as phonemes). If the graphemes, modified with graphons, represent CV syllables, then graphons can be interpreted as variants of vowels, even being tied closely to the shape of a letter. The difference between this script and the alphabetic type lies in the nature of these vowel marks in an akshara: they are not the full-formed graphemes but only graphons, though they represent a necessary component of a syllable, the 'soul' of the akshara as traditional Tamil descriptions say. This is, thus, a transitional type of a system, which combines features of alphabet and abugida.

Yet, the later Tamil Brahmi restores the syllabic mode of writing and differentiates it further: it adds a new operator: a mark of vowel rejection. This mark works

like a diacritic *virama*, or *halanta*, that can be found in later North Indian scripts. It enlarges the paradigm of a syllabic sign, thus confirming the abugida principle.

	a	m	ma	mā
Ashokan Brahmi (III BC)	𑀓	𑀣	𑀤	𑀥
Early Tamil Brahmi (III BC)	𑀓	𑀣	𑀤	𑀥𑀓
Later Tamil Brahmi (II AC)	𑀓	𑀣	𑀤	𑀥
Modern Tamil	அ	ம	மா	மா

Figure 3. Marks for a vowel /a/, a pure consonant /m/, and for vowel differentiated consonants /ma/ and /ma:/¹²

So a vowel in the abugida script may be represented by a full-sized grapheme only in initial or independent position (after another vowel).

4.4 Tibetan

In many derivative writing systems, the vowels get their own shapes in independent positions. Still some scripts, like Kharosthi, prefer the empty-consonant device in representing independent vowels. Tibetan writing forms graphemes for other independent vowels with diacritics on the base of the *a*-grapheme; yet, there are two vowels included in the alphabet for representing /a/ ([hà/?á] and [hà/?à]): the first one (the empty-consonant) is the base for short vowels, while the second can serve as a diacritic for marking long vowels (it is more correct to say that the postscript <’a> acts de jure as a ligatured consonant: it forms a base of an akshara with a long vowel – ‘prolonged akshara’ – and vowel diacritics are added to this compound akshara).

ཨ	<a>	[hà/?á]	ཨ	<’a>	[hà/?à]
ཨི	ཨུ	ཨེ	ཨོ	ཨཱ	ཨི
i	u	e	o	ā	ī

Figure 4. Some Tibetan independent vowels

4.5 Thai

In Thai script, initial vowels also exploit the base of a grapheme *o ang* included in the alphabet. The great variety of vowels in the Thai language uses almost

all possible positions for vowel marks around the consonant grapheme, and the same marks are used around the *o ang* for independent vowels. Here it may be an efficient device. So in Thai writing there are not only *sub-* and *superscripts*, *pre-* and *postscripts*, but also *circumscripts* and even more complex designations.

In some other derivative systems, an analogous compromising way of initial vowel designation can be found in the use of an empty consonant as the base for a standard akshara; but in fact it is just a sign from the alphabet. An empty consonant is an unreadable sign; it serves as a carrier for an independent vowel and thus maintains the structural uniformity of graphemes. Such a device is used in Rong-ring (Lepcha) script, derived from Tibetan. A comparable way of writing initial vowels after the unreadable consonant can be seen in Korean script which, though not a descendant of Brahmi, uses this device for balancing syllable representation. Yet in Korean, the consonant /ng/ is unreadable only in the initial position of a syllable block marking the initial vowel; the same consonant is read in the final position. So it cannot be regarded as using the Tibetan pattern.

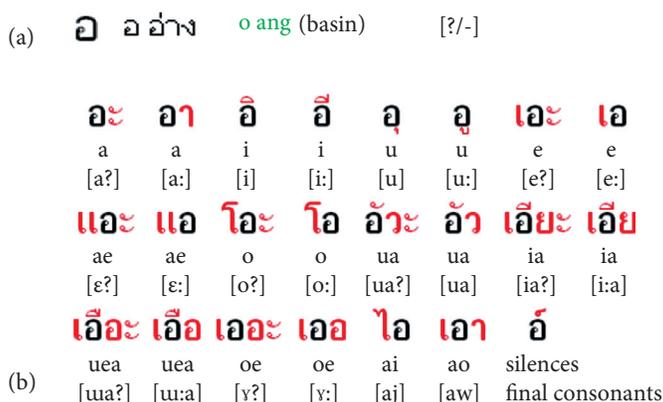


Figure 5. (a) Thai basic vowel grapheme /o ang/ with its mnemonic name ‘basin’; (b) some vowel graphemes¹³

The Brahmi model of using different independent shapes for initial vowels was also widely applied. In Khmer script, for example, most vowels are represented by different shapes in initial position. But there is also an economic way of vowel modification in consonant aksharas: two series of consonant aksharas (originally voiced and voiceless) have identical consonant meaning in pairs, but the inherent vowel differs – /a/ or /o/, and the identical vowel marks (there are about 30 vowels in Khmer) have different meanings depending on the consonant series they refer to.

4.6 Diacritics vs. graphons

In a dependent position (after a consonant), a vowel can be marked by a graphon closely tied to the main sign (radical), thus forming an integral sign; this device, used in early writing systems,¹⁴ was later developed by separating two parts: a dependent vowel symbol, standing distantly from the main sign, and the main sign itself, which remains unchangeable. Such way of writing allows the main grapheme to be conserved intact in its own space in line and also makes the marks for the same vowel uniform (for they need not adapt to the shape of the main graphemes, cf. Ku: Lu in Brahmi, Figure 2b). These symbols differ from graphons and may be regarded as stable alienable marks of their phonological function, i.e. as pure diacritics, *matras*. This makes the script more distinctive, as we can see in Devanagari or in the Tibetan script.

Now two types of the vowel marks can be distinguished: (1) those closely tied with a main grapheme within its inner space (graphons); and (2) those bound to the space of a main grapheme from outside (diacritics). Their common characteristics are their exclusively dependent use and their size, which is usually smaller than that of a main grapheme.¹⁵ Graphons and diacritics differ in their relative position (inside or outside the inner space of the main grapheme) and in their mode of conjunction to the main grapheme, solid or distinctly separable. This behavior resembles fusion or agglutination in morphology. Therefore, I consider a diacritic as a writing symbol which can modify a grapheme; its subordinate phonological function – to distinguish between variants of an invariant unit – determines its position (usually outside the inner space of a grapheme) and size (usually not equal to that of a grapheme horizontally and /or vertically); its mode of conjunction with a grapheme is distinct. Full-sized representations of vowels placed in line but different from their initial forms (such as for example marks for /e/ in Bengali, Tamil, and others) represent quasi-graphemes for they fulfill the subordinate function of diacritics, the vocal modification of aksharas, which in this way become complex.

In a way, there is a parallel between akshara-formation and word-formation (word morphology): graphons and diacritics function like internal inflexion and affixation, respectively:

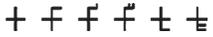
SING – SANG – SONG	PLAY – PLAYED – DISPLAY
	
ka k̄a ki k̄i ku k̄u	ka k̄a ki k̄i ku k̄u

Figure 6. Comparison of devices in word morphology and akshara-formation

4.7 Agglutination vs. fusion: Tamil script

A more complex system of vowel representation can be found in Modern Tamil. Besides the set of initial signs and the set of regular diacritics, there is a number of non-standard aksharas where a consonant gets its vowel modifications in an irregular way. Thus, an akshara can have the shape of a free modification or of a ligature in which consonant and vowel shapes are combined in mutual fusion. Such inseparable shapes tend to function as Kana signs (cf. Sproat 2006:57). Yet, in comparison with the invariant form, a common base can be found; moreover, vowel marks can have somehow distributed alloforms (three non-standard forms for /-u/), though the same form can refer to different vowels (cf /-a:/ and /-u/).

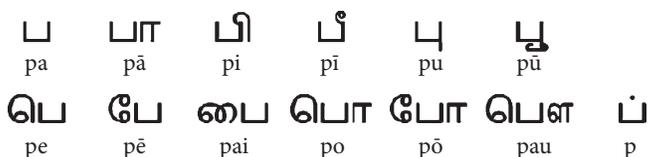


Figure 7a. Regular vowel marks



Figure 7b. Some irregular vowel marks



Figure 7c. Alloforms for /-a:/ and /-u/¹⁶

So, Tamil aksharas can be opposed in the mode of conjunction: agglutination or fusion. Thus, the paradigmatic system of Tamil script has four main components: vowel aksharas, Ca (consonant) aksharas, a matrix of regular vowel modifications, and a set of non-standard vowel modifications.

5. Opposition of short and long vowels

5.1 Devanagari

The use of diacritics is usually limited to the outside positions above and below the line, so they are divided into *superscripts* and *subscripts*. However, in Devanagari there are also complex marks for dependent vowels: a full-letter high stem (vertical right stroke) as a mark for the long phoneme /a:/ and as a base for superscript vowel diacritics. (In some way it resembles the Arabic *alif* in its function of a vowel carrier.) The stem cannot be regarded as a proper grapheme because it does not constitute a full letter, but when added to the main sign, it extends the letter space in the horizontal dimension. This is an appropriate, iconic way for representing long or complex sounds, i.e. long vowels and diphthongs (/a:/, /i:/, /o/, /au/). Devanagari uses this stem also for the short /i/, in this case placing it with upper diacritic before the main grapheme; such inversion cancels the operation of lengthening. So according to its position, the vertical stroke can be a *postscript* or a *prescript*.

There is no reason to consider the vertical stroke as an empty consonant, since it does not form a base for each vowel, and moreover, it is also an inalienable element of many consonant syllable signs, i.e. a carrier of a distinctive trait for 24 of the 36 letters of the script and for a quantity of vertical ligatures. It also creates the vertical dimension of a sign and determines its boundary. Thus, it helps to divide the letters from each other and it maintains the uniformity of the style of writing together with the horizontal stroke binding letters. In these functions, the vertical stroke can be compared to a round contour, an outline common to many signs in Oriya script (due to the historical material of writing: palm leaves that were torn by horizontal lines). In quite a different system in another part of the world, in Maya script, the element of this kind is an outside line of many pictographic glyphs. It has the shape of a face (or just an oval), in which distinctive features are placed.

ସାନ ଭାଣ୍ଡା ନାମ କନକମଞ୍ଜରୀ

Figure 8. Oriya script

5.2 Gurmukhi

Thus, the vertical right stroke may be an alienable or unalienable sign. As an alienable sign, it has two functions in Devanagari: it marks the long /a:/ or it serves as a carrier for a vowel diacritic. The first function is fulfilled by other vowel diacritics too, but the size and position in line discerns the stroke from them. Gurmukhi

writing differentiates further functions: it has a shorter stroke for /a:/ (turning it to an *upper postscript* – *kannā*), while for the vowel carrier it reserves the full size (moreover, there are not so many consonants with a stroke in Gurmukhi). In such a way, the long stroke functions only as a carrier (as a part of a sign) and the shortened stroke becomes a postscript diacritic marking a long /a:/. The other pairs of short and long vowels (i, u) are marked similar to those in Devanagari.

Thus, vowel marks include postscripts and prescripts, which are not full-size letters in the horizontal dimension. It is remarkable that there are three different graphic shapes (bases for /a/, /i/, /u/) for 10 initial vowels (divided in groups: back, front, and low), but only one of them (for /a/ or schwa) can really be used independently; the others obtain their different value with diacritics, common to vowel designations for consonants, and can therefore be regarded as vowel-type carriers. Thus, Gurmukhi script has a compromising way to designate vowels. Bhatia (2003: 181–213) regards this mode as a device of generalization, which Devanagari failed to capture. Yet, Gurmukhi script is a more recent development (16th century).

The relevance of writing style is based on the balanced proportions of graphemes and ‘sub-graphemes’ of dependent vowels.

ਸਾਰੇ ਇਨਸਾਨ ਆਜ਼ਾਦ ਅਤੇ ਹੱਕ ਤੇ ਇੱਜ਼ਤ ਦੇ ਲਿਹਾਜ਼ ਨਾਲ
ਬਰਾਬਰ ਪੈਦਾ ਹੁੰਦੇ ਹਨ। ਉਹ ਅਕਾਲ, ਸਮਝ ਤੇ ਚੰਗੇ ਮੰਦੇ ਦੀ

Transliteration: *sa:re insa:n a:za:d ate hakk te izzat de liha:z na:lbara:bar paida: hu'de han* ⟨...⟩

Translation: ‘All human beings are born free and equal in dignity and rights.’ ⟨...⟩

(First sentence of Article 1 of the Universal Declaration of Human Rights)

Figure 9. Gurmukhi script; transliteration of the first sentence (till the vertical stroke – danda as a point mark); translation

The most consistent representation of long vowels can be found in Tibetan script where a unique subscript diacritic serves to indicate a long vowel (both independent and in Ca-akshara) marked with another standard diacritic (Figure 4).

6. Is there any progress in vowel representation?

While the main opposition of high and low vowels /i/ vs. /u/ is usually represented with super- and subscripts in many Brahmi derived scripts, there is an increasing number of vowel diacritics in line. Bengali script chooses a prescript position marking a vowel /e/ with its own diacritic; there are many examples of in line vowel diacritics in Thai, Khmer, and Tamil writings. They can have full size

and when composed of two parts (pre- and postscript in circumscripts) they take up more space than the main sign, so an akshara can be represented by a complex of signs: grapheme and quasi-graphemes. Sometimes they can form ligatures with main signs and behave like consonant signs and, as a result, the script becomes oriented in a more linear way (though there may be some non-standard ligatures). However, they remain diacritics as they depend on main (consonant) signs and have different initial forms. The vocalization can therefore be regarded as a phonological feature of a syllabeme generating its variants, the proper function of diacritics.

Some gradual changes in the representation of vowels in writing systems can be seen, moving from abjad to alphabet. While in abjad vowels are concealed under the shape of a consonant in a way that there is no vowel differentiation (if it is not specially emphasized), abugida systems allow consonants to take different shapes 'clothed' by diacritics of different vowels.¹⁷ Ultimately, alphabetic systems represent vowels by letters equivalent in size and position to consonant letters. The Greek 'claim to democracy' gives them their independent status in the Greek alphabetic script and derived scripts.

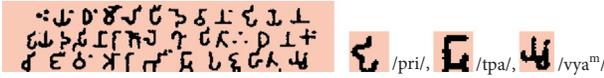
Of course, this is only a metaphor; Greeks succeeded in alphabetic writing due to the analytic approach applied to their language structure (and maybe due to the inaudibility of Semitic *alif* and *ain* for their phonematic hearing). Abugida scripts are no less exact in representing sound sequences; and the degree of exactness is relative even in modern alphabets. The indefinite vocalization in the abjad is more appropriate for maintaining the recognizability of radicals. So the question of progress is more a rhetorical than an evaluative one.

7. The representation of consonant clusters: Ligatures

7.1 Brahmi

Brahmi script has a model for constructing ligatures, or conjunct consonants (*samyuktākshara*), to represent consonant clusters. Brahmi does not have many ligatures: for /pra/, /tpa/, /vya/ and some others (mostly with a sonorant).¹⁸ The model of conjunction is vertical, the base could be chosen as the best graphic variant (e.g. in the Ashoka's Rock edict I, in /pra/ *r* is rendered by a superscript, in /tpa/ *t* is rendered as a subscript, in /vya/ *y* is the base and *v* is the subscript). The whole conjunct tends towards a minimal extension of its space and exploits not the linear principle of connection in succession but the emblematic principle of a balanced graphic composition. The ligature does not differ from the vowel differentiated

akshara: it forms an integral sign in a line. The proper diacritic that does not touch the main sign is an *anusvara*: a right superscript dot for nasalization.



Transliteration: iya^m dha^mmālipī devāna^mpriyena // priyadasinā rāñā lekhatā i dha na ka^m // ci jīva^m ārabhi^mtpā prajūhitavya^m (Schneider 1978)

Translation: ‘Beloved-of-the-Gods, King Piyadasi, has caused this Dhamma edict to be written. Here (in my domain), no living beings (are to be slaughtered or offered in sacrifice...)¹⁹

Figure 10. Ashoka’s Rock Edict I (Girnar 257 BC), a fragment with aksharas: /pri/ (the third from the right, in the first line and the first in the second line), /tpā/ (the sixth from the right, third line), /vya^m/ (the first from the right, third line)

7.2 Devanagari

The technique of combining consonant signs into ligatures was developed in derivative scripts. In Devanagari, two types of ligatures exist: horizontal and vertical. Vertical ligatures have a fixed order from top to bottom. Though a conjunct presupposes a common space for its components, they are distinctive and only have one common vertical stroke (if any), which is not longer than the line; that is why the vertical ligature is usually compressed. The connection is systematic, agglutinative, and subordinate to the rules (yet there are exceptions when one of the elements (or both) is modified in fusion). Thus, a consonant ligature differs from a vowelled akshara as a unique integrated sign within the line, while an akshara with a vowel diacritic has two discontinuous parts, one in line and the other (usually) out of line. The text, consequently, has two dimensions: horizontal, as a chain of aksharas in line, and vertical, along a stroke, in a stack – within the akshara. Though the dimensions of a ligature can be enlarged (like in aksharas with post- or prescript vowel diacritics), it represents a unique sign, as can be seen just in a simple example of writing numerals (with a horizontal ligature for /nya/ and vertical ligatures for /dv-i/, /tr-i/, /ñc-a/, where /tr-i/ represents modified forms of components in fusion):

०	१	२	३	४	५
शून्य	एक	द्वि	त्रि	चतुर	पञ्चन्
śūnya	eka	dvi	tri	catur	pañcan
0	1	2	3	4	5

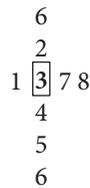
Figure 11. Some Devanagari numerals

Vertical outline space is usually used for vowel designation in aksharas, while horizontal extension is common for ligatures as well as for the addition of a vowel element (the stroke) in an akshara.

7.3 Tibetan

The rules for conjunction in Tibetan are more complicated, as sound sequences are not so balanced and the structure of a syllable can be (CC)CV(CC) with a maximum of three consonants before the vowel and two consonants after it. As a consequence, the Tibetan script has a two-dimensional way of representing consonant clusters within a syllable: (1) the vertical stacks of graphemes can be completed with one to two graphemes in line; (2) the conjunct elements may be placed before and after, above or/ and under the central one, which is the 'root' of a syllable. Stacks of signs usually 'hang' down from the upper horizontal line of a grapheme (which is named 'head' and can be regarded as a sign-carrier).

Therefore, the order of reading may be as follows (the numbers mark the order of reading):²⁰



A root grapheme is situated in the center (3). A syllable could be represented by four graphemes in horizontal dimension, in series standing separately. A vertical stack can contain four to five components: a main grapheme (root), a consonant superscript, one to two consonant subscripts, and the vowel diacritic (subscript or superscript); the vertical compression (as in Devanagari) is not adopted, so the vertical dimension creates compound graphic signs in which subscript consonants are common.

As a complex akshara may include more than one sign in line, a problem of dividing closed syllables arises. It is solved by introducing a dividing dot in the upper right side of the last sign of a syllable, a 'hanging dot' called *tsec*. The last sign in a group before the dot (and without vowel diacritics) receives a 'trimmed' sound meaning C – without a vowel. So, a grapheme can have a syllabic or an alphabetic meaning according to its position: only the root is a syllable, but conjunct consonants in line can be regarded as graphemes with reduced phonological meaning. Thus, a closed syllable can be represented as a complex of graphemes (a complex akshara).

༠	༡	༢	༣	༤
ལྷན་ཀོར་	གཅིག་	གཉིས་	གསུམ་	བཞི་
laykor	chig	nyi	sum	shi
0	1	2	3	4
klad-kor	gchig	gñis	gsum	bzhi

Figure 12. Some Tibetan numerals with transcription and transliteration²¹

8. Analytic vs. synthetic writing

Simple graphemes – consisting of one ‘radical’ for Ca – can be distinguished from compound ones, burdened with subscripts or superscripts, or bound with ligatures. It is possible to draw a parallel between this difference and the analytic/ synthetic character of word sequences in a grammar (in accordance with the typological approach to language analysis by Greenberg (1960: 178–194)): analytic chains do not integrate radicals and affixes, while synthetic chains integrate them. Similarly, analytic writing presupposes the use of simple graphemes, corresponding to a standard pronouncing unit (a sound/ a mora/ a syllable), forming a standard phonological unit (a phoneme, a ‘moreme’, a syllabeme),²² while synthetic writing presupposes the use of compound graphemes. The opposition of simple and compound graphemes allows speaking about the degree of synthesis (complexity) of a grapheme (S_g) and the degree of synthesis (complexity) of a writing system (S_w), which can be calculated.

The extension of the vertical dimension used as a graphic variant in Brahmi for representing vowels and ligatures became more frequent in Devanagari (though vertical ligatures are compressed in line) and quite common in Tibetan. The further development of writing may have exploited this vertical dimension as a prevalent tendency: it exists in Phags-pa script, which is derived from Tibetan but culturally oriented to Chinese script in its appearance. The Phags-pa script deviates completely from two-dimensional writing and changes the direction of writing from horizontal to vertical (in a top to bottom fashion). Vowel diacritics and conjunct consonants are written in a common stack for a syllable; all components of a syllable are bound together in a linear block. Thus, the Phags-pa script seems closer to alphabetic writing, yet the structure of abugida is preserved in two main points:

- a. there is no mark for /a/ in the dependent position (after a consonant), therefore this ‘consonant’ grapheme corresponds to a Ca syllable;
- b. shapes of vowels differ in the independent (initial) and dependent position.



Figure 13. Phags-pa script

The Phags-pa script led the way to the new integration of signs: putting them in a syllable block where previous (Tibetan) vertical diacritics are no more opposed to previous in line signs, since they all are arranged in a stack. As a result, the problem of delimitation arises for a model of conjunction can be fusion rather than distinct agglutination. So Phags-pa reveals great complexity in decoding its blocks, and the problem of graphic complexity is not merely reduced to the degree of synthesis.

A compromising way is used in Khmer, for instance. There the consonant ligatures are extended downwards by the subscript consonants while vowel marks can have different positions around the ‘root’ grapheme (similar to what is found in Thai where vowels can also form non-standard ligatures); the ‘consonant’-dimension becomes downwards oriented. Subscript consonants can have a compressed or modified form. As a result, the paradigmatic system can have an additional component of consonant ligatures.

Another way of consonant-vowel arrangement in a syllable block was made use of in Hangul, the alphabetic Korean script (not derived from Phags-pa): though a syllable cannot begin with a vowel, the orientation of a consonant grapheme is subordinate to the graphic shape of a vowel grapheme; thus, the allograph forms appear in the block construction.

While speaking about graphic complexity of writing, different parameters should be taken into account: the degree of synthesis, the mode of conjunction (fusion or agglutination), the general number of graphemes and sub-graphemes in a system, the number of allographs and of non-standard ligatures, and also the presence of word (syllable) division marks. Finally, writing styles (cursive, calligraphic, or printed) can differ greatly in these respects. In addition, graphic complexity does not correspond to orthographic complexity which is connected mostly with language change (cf. Figure12).

9. Tone differentiation

The development of related writing systems tends to express speech characteristics more distinctively. The necessity of marking different tones produces not only a functional specification of graphemes (in Gurmukhī), but also new series of diacritics, as in Thai script or in Burmese script. So diacritics can have different functions there.

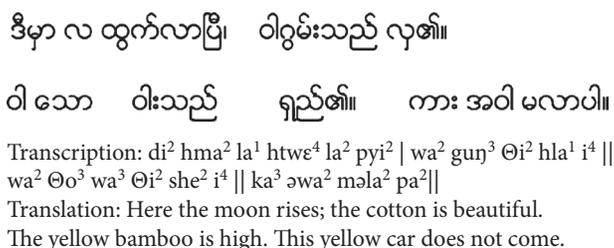


Figure 14. Burmese script

Some conjunct consonants can be represented by subscript diacritics. Vowels have prescript, postscript, superscript, or subscript marks and their combinations; they serve at the same time to specify the tone, which is represented in a non-standard model (differently for different vowels); tone marks are “intricately bound with vowel marks” (Daniels 2006: 19). Burmese script, consequently, has a very complex system of graphic representation, as well as orthographic difficulties.

10. Text division

Text division was not regarded as being necessary for a long time. In Brahmi script, marks of sentence division are rare. In Devanagari, the mark is a horizontal bar that binds the graphemes in words or groups of words (*matrica*, proceeded from Kushan Brahmi). There is also a mark for the end of a sentence, a vertical right stroke higher than the line (*danda*). All these signs, which should be considered as punctuation marks, make the script more distinctive. Later they developed into a more complicated system.

The most interesting punctuation marks are used in the Javanese script. Unreadable symbols at the beginning of a letter can designate an addressee’s social status by the height of the left line. Some honorific graphemes (*akshara murda*) are also used for writing a high ranking person’s name.



Figure 15. Honorific marks at the beginning of Javanese letter

There is a certain parallel with honorific writing in the Russian epistolary tradition, where an author capitalizes the first letter in the pronoun (and its grammatical forms) in addressing a person of high status, while choosing the lowercase letter when he/she does not want to mark his/her high respect or distance, cf. *Bbl* vs. *bbL*. (So я 'I' is always in minuscule while 'you' can have three forms: *mbL*, *bbL*, *Bbl*).

11. Calculations of the degree of synthesis

Simple or complex aksharas can determine the analytic or synthetic writing. The index of synthesis of writing (S – degree of synthesis) can be calculated, which I will briefly demonstrate in the following illustration.

As an example, the writing of the famous mantra *Om mani padme hum* can be compared in the following scripts:

Devanagari:	ॐ मणिपद्मे हूँ	\bar{U}^m ma-ṇi pa-dme hu ^m
Bengali:	ওঁ মণিপদ্মে হুঁ	O^m ma-ni pa-dme hu ^m
Tibetan:	ཨོཾ་མ་ཎི་པདྨེ་ཧཱུྃ།	O^m ma-ni pa-dme hū ^m
Tamil:	ஓம் மணி பத்மேஹும்	\bar{O} -m ma-ni pa-d -me hū-m

The transcriptions are divided in graphic syllables (aksharas), and their number A is: $A^D = 6$, $A^B = 6$, $A^{Ti} = 6$, $A^{Ta} = 9$. The number G of elementary phonologically meaningful (or distinguishing phonological meanings) graphic elements is: $G^D = 13$, $G^B = 12$, $G^{Ti} = 14$, $G^{Ta} = 16$. (Yet there exists another possibility to write *-dme-* in Devanagari: in two aksharas with a visarga after *-d-* (*d-me*) like in Tamil; in this case $A^D = 7$, $G^D = 14$, $S^D = 2$. The same is true for Bengali.)

Some calculations need comments:

1. The unique graphic element for vowel prolongation in Devanagari (U: \bar{U}) and in Tamil (O: \bar{O}) is differentiated, using the procedure of Greenberg's square.
2. A vertical stroke is not divided from a superscript in vowel diacritics (for /i/) in Devanagari, in Bengali, and in Tamil because a sole differentiating meaning

for the stroke cannot be determined (yet this can be done otherwise when considering two homographic signs: one for a vowel carrier and another as an unalienable part of many letters).

3. The inline sign for /e/ in Bengali and in Tamil are considered elementary graphic elements, a diacritic (or quasi-grapheme, but not a true akshara-grapheme).
4. The close ligatures for /dm-/ in Devanagary and Bengali are counted as two graphic elements, as they can be recognized. (This may be arguable since the model of conjunction is fusion rather than agglutination (cf. Tibetan), but it can be specified in introducing the index of fusion.)
5. The vowel carrier (grapheme <a> [há/?á]) in Tibetan is considered a meaningful graphic element, but an empty sign.
6. Sentence marks and syllable dots in Tibetan are not counted.

On this base, the degree of synthesis S can be calculated as a ratio of G to A:

$$S^D = 13:6 = 2.17$$

$$S^B = 12:6 = 2.0$$

$$S^{Ti} = 14:6 = 2.33$$

$$S^{Ta} = 16:9 = 1.78$$

Tamil yields the smallest result, which is clear since it uses *pulli* (virama) and does not have consonant ligatures. Sproat (2006: 56) refers to the formal point of view regarding Tamil script as the simplest Indian writing.²³ The Tibetan writing is the most complex in regard to synthesis. Of course, these results are quite preliminary and cannot be considered as definitive indexes of synthesis for these writing systems. The reason is, above all, the fact that very short phrases were compared. The examples just demonstrate the possible method of index calculating. The degree of synthesis is not the unique criterion for evaluating the complexity of a script; other dimensions noted earlier should be taken into account.

12. Conclusion

In conclusion, it should be stated that the model of abugida, once invented in India, has produced many variants in the related scripts of different languages. The inherent possibilities have developed in further differentiation. This is a common way in the adaptation of writing: through more distinctive differentiation to new integrations. In this text, I tried to show how derivative scripts maintained and developed structural oppositions in 'akshara grammar': in the representation of a phonological unit as a non-linear composition – a 'written emblem'. Simpler

writing tends to correspond to the linearity, while more complex writing tends to exploit the emblematic principle as well.

The productive use of emblematic syllable representation is characteristic also for the Korean script which has an alphabetic nature, but is not quite simple. Its difficulty stems from the emblematic arrangement of its graphemes in syllable blocks and their changing shapes (allographs), which need to be discovered in reading. An emblem tends to represent the whole in integrating its parts in a conventional way.

It can be assumed that the linear and emblematic principles of writing provide productive possibilities that can be exploited constructively. Sometimes a linguistic emblem, an abbreviation, can take the form of a written emblem, so both of them can have a symbolic meaning.



Notes

1. See e.g. Bright (2000:63–71) and the recent discussion of Daniels (2009:277–281) and Swank (2009:282–285).
2. The standardization of the logosyllabic Yi writing for the Nosu language in the 1980s resulted in a syllabic system of 819 basic characters (Bradley 2009:179).
3. The name *moraic* for this type (Sproat includes also Linear B, Sumerian and Mayan writing systems in it (cf. Rogers 2005:274)) seems not quite appropriate since it implies that graphemes encode moras (syllables with a short vowel) only in this type of writing and not in others like abugida or abjad which remains debatable.
4. Thus, an alphabet can have properties not only of a system, but also of a text that can be learned and interpreted, and its format can be used with other purposes (semiotic – counting – and even magic).
5. In fact, their names can be partly derived from their pronunciation as in Latin or modern Cyrillic alphabets, yet they become not only marks of sounds, but names of letters (they differ from hieroglyphs which can render sounds or words or morphemes).
6. References can be made to: Lyavdansky A.K. (2009). The origin and early development of the West Semitic alphabets. In A. Belova, L. Kogan, S. Loesov & O. Romanova (eds), *Languages of the World. The Semitic languages. Akkadian. Northwest Semitic*. Moscow: Academia, 811–821.
7. The problem of the origin of Indian scripts is reviewed in Salomon (1995); mostly, Kharoshthi is seen as an invention under the influence of Aramaic script.

8. There are some CCV graphemes, too (<ksa>, <sta>), which are elementary shapes and not ligatures.
9. Fragments are extracted from the illustrations in: <http://www.rbardalzo.narod.ru/novosti.htm>.
10. This can be found only in later documents, maybe under the influence of Brahmi, as Salomon (1996: 373–383) supposes.
11. Fragments are extracted from the illustrations in: <http://www.omniglot.com/writing/thai.htm>.
12. Fragments are extracted from the illustrations in: http://en.wikipedia.org/wiki/Tamil_Brahmi with reference to: Mahadevan, Iravatham (2003). *Early Tamil epigraphy from the earliest times to the sixth century A.D. (Harvard Oriental Series 62)* Cambridge: Harvard University Press, 173.
13. Fragments are extracted from the illustrations in: <http://www.omniglot.com/writing/thai.htm>.
14. The device of modification of a sign in its inner space is still used in more recent writing systems on the base of Latin or Cyrillic writing; like other devices – superscript and subscript marks, digraphs, ligatures – they usually just create another alphabetic sign and do not operate like true graphons or diacritics whose function is to mark regular changes of sounds. Daniels considers such marks with the example of Vietnamese vowels as “integral parts of the letters” (Daniels 2006: 19). Such marks can be named quasi-graphons or quasi-diacritics, if needed.
15. This is not the case in South and South-Eastern scripts with a quantity of vowels.
16. Fragments are extracted from illustrations in: <http://www.omniglot.com/writing/tamil.htm>.
17. The metaphor of ‘clothes’ is used for the vowel diacritics in Javanese script.
18. Sound sequences in Pali have less consonant clusters than in Sanskrit or some later Indian languages like Hindi (Shevoroshkin 2004: 135–138; the author of the book which was firstly edited in 1969 and was not translated into English proposes two ways of translating his term from Russian: ‘sound chains’ or ‘sound sequences’).
19. Translation according to Dhammika, Ven. S. (1993). *The edicts of King Ashoka*. Electronic edition.
20. The scheme is based on Rogers (2005:225), using more detailed description in: http://tibetan.bitecs.ru/files/text_intros.pdf.
21. Transliterations given according to: http://tibetan.bitecs.ru/files/text_intros.pdf.
22. While linguists have been elaborating the notions of phoneme, morpheme, and syllabeme during the last 100 years, creators of writing systems resolved the problem of generalization of pronouncing units in graphemes thousands of years ago. Their analysis can be considered as a practical base for further investigations. So the notion of ‘moreme’ can be put in a line with others according to the logic of graphic systems using them.
23. Still he shows that non-standard vowel-consonant ligatures can resist analysis, which makes Tamil move closer to Japanese Kana, a ‘core syllabary’ (Sproat 2006: 56–68); it can enlarge the general number of graphemes that resist analysis (and the general complexity of writing system), but not the degree of synthesis.

References

- Bhatia, Tej K. (2003). The Gurmukhi script and other writing systems of Punjab: History, structure and identity. In Peri Bhaskararao (ed.), *Working papers of the international symposium on Indic scripts: past and future*, 181–213. Tokyo: ILCAA.
- Bradley, David (2009). Language policy for China minorities: Orthography development for the Yi. *Written Language and Literacy* 12: 170–187.
- Bright, William (1996). The Devanagari script. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*, 384–391. Oxford: Oxford University Press.
- Bright, William (2000). A matter of typology: Alphasyllabaries and abugidas. *Study in the Linguistic Sciences* 30: 63–71.
- Coulmas, Florian (1996). *The Blackwell encyclopedia of writing systems*. Oxford: Blackwell.
- Daniels, Peter T. (1996). The study of writing systems. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*, 3–17. Oxford: Oxford University Press.
- Daniels, Peter T. (2006). On beyond alphabets. *Written Language and Literacy* 9: 7–24.
- Daniels, Peter T. (2009). Two notes on terminology. *Written Language and Literacy* 12: 277–281.
- Fedorova, Liudmila (2011). К построению типологии письменных систем (Towards the typology of writing systems). *Moscow Journal of Linguistics* 13: 174–195 (in Russian).
- Gamkrelidze, Tamaz V. (1994) (1989). *Alphabetic writing and the old Georgian script*. New York: Caravan Books.
- Greenberg, Joseph H. (1960). A quantitative approach to the morphological typology of language. *International Journal of American Linguistics* 26: 178–194.
- Hymes, Dell (1974). *Foundations in sociolinguistics*. Philadelphia: University of Pennsylvania Press.
- Ivanov, Vyacheslav V. (2004). *The linguistics of the third millennium: Questions to the future*. (In Russian) Moscow: Iazyki Slavianskoi Kul'tury.
- Lyavdansky, Alexej K. (2009). The origin and early development of the West Semitic alphabets. In Anna G. Belova, Leonid E. Kogan, Sergej V. Loesov & Olga I. Romanova (eds.), *Languages of the world: the Semitic languages. Akkadian. Northwest Semitic*, 811–821. Moscow: Academia.
- Neef, Martin (2010). On defining phonology. In Sang-Oak Lee (ed.), *Contemporary Korean linguistics: International perspectives*, 216–230. Gyooha-eup (Korea): Thaeaksa.
- Pandey, Pramod (2003). Phonetic and phonological bases of Hindy orthography. In Peri Bhaskararao (ed.), *Working papers of the international symposium on Indic scripts: past and future*, 41–61. Tokyo: ILCAA.
- Rogers, Henry (2005). *Writing systems: A linguistic approach*. Malden, MA: Blackwell.
- Salomon, Richard G. (1995). On the origin of the early Indian scripts: A review. *Journal of the American Oriental Society* 115: 271–279.
- Salomon, Richard G. (1996). South Asian writing systems: Brahmi and Kharoshthi. In Peter T. Daniels & William Bright (eds.), *The world's writing systems*, 373–383. Oxford: Oxford University Press.
- Schneider, Ulrich (1978). *Die großen Felsen-Edikte Ásokas. Kritische Ausgabe, Übersetzung und Analyse der Texte*. Wiesbaden: Otto Harrassowitz.
- Shevoroshkin, Vitaly V. (2004) (1969). *The sound chains in the languages of the world*. (In Russian) Moscow: URSS.

- Sproat, Richard (2006). Brahmi-derived scripts, script layout, and segmental awareness. *Written Language and Literacy* 9: 45–66.
- Swank, Heidi (2009). A response to Peter T. Daniels. *Written Language and Literacy* 12: 282–285.
- Wali, Aamir, Richard Sproat, Prakash Padakannaya & Bhuvana Bhuvaneshvari (2009). Model for phonemic awareness in readers of Indian script. *Written Language and Literacy* 12: 161–169.

Author's address:

Liudmila L. Fedorova
Russian State University for the Humanities
Institute of Linguistics
Miuskaya square 6
Moscow 125 993
Russia

lfvoux@yandex.ru