

DOI: 10.2478/linpo-2013-0013

THE DEVELOPMENT OF GRAPHIC REPRESENTATION IN ABUGIDA WRITING: THE AKSHARA'S GRAMMAR

LIUDMILA L. FEDOROVA

ABSTRACT: Liudmila L. Fedorova. *The Development of Graphic Representation in Abugida Writing: The Akshara's Grammar*. The Poznań Society for the Advancement of the Arts and Sciences. PL ISSN 0079-4740, ISBN 978-83-7654-274-4, pp. 49–66.

Phonological writing systems can use different modes in the arrangement of phonological information: linear or emblematic. The latter presupposes a two-dimensional composition of graphic elements which convey information of different levels: basic phonological units or their secondary features. Abugida is one of the most perfect modes of writing, using a two-dimensional principle of organization in its graphemes. Indian scripts imply forms of abugida, based on aksharas – orthographic syllables – as main graphemes, with their vocalization marked by sub-graphemes. Their complex forms reveal hierarchical structures representing phonological structures. Thus, “the akshara’s grammar” may be described, with its paradigms of shapes and meanings, their organization in space, their modes of conjunction. The functional characteristics of diacritic types and ligatures can be specified. The graphic shapes of the akshara can be represented as forming a “grammar” comparable to morphological structures. The graphic complexity of writing can then be estimated in typological perspective. This paper is an attempt to describe such a grammar.

KEY WORDS: Writing systems, abugida, Indian scripts, akshara-formation, complexity of writing

Liudmila Fedorova, Russian State University for the Humanities, Moscow, lfvoux@yandex.ru

1. INTRODUCTION

The development of writing systems reveals the interaction of different forces and tendencies: tradition and innovation, redundancy and laconism. Simplicity or complexity of writing is a result of the interaction between these. However, these qualities are relative and require estimation. The tendency to render phonological oppositions of the language more distinctly follows the interests of the reader; this tendency is governed by a demand for exactness, which in turn may lead to complexity of graphemic structures. This tendency is counteracted by another one, the tendency of the writer to save effort (demand for economy or just laziness), which may lead to simplification of graphemic structure.

This paper aims to demonstrate, based on the example of akshara shapes, how the demand for exactness serves to elaborate more complex graphic forms in creating a writing system, or in the adaptation of a borrowed script to new language structures. In so far as a writing system is a graphic representation of a language, it tends to follow language chang-

es; yet the force of tradition may restrain graphic changes, creating orthographic difficulties. The problem of the complexity of writing and its estimation is therefore in question.

The structures of an akshara constitute paradigms that can be compared with morphological structures. Some phenomena reveal similarities and may be described in parallel terms. These problems will be examined based on the data of Indian scripts.

2. SOME DEFINITIONS

Akshara is the traditional name for the main graphic sign in the Indian scripts; it usually corresponds to a pronunciation unit – a syllable; thus, an akshara is an orthographic syllable.

The Indian scripts, or more exactly, various writing systems in South and South-East Asia, are descendants of the Brahmi script (which is recorded since the 3rd century BC). This script (like the contemporaneous Kharosthi script) differs from the Aramaic and Greek writing systems in the way of representing sounds (phonetic units). Brahmi is more exact than the Aramaic script in reflecting vowels and it is more compact than the Greek script in its graphic realization (since only one grapheme is used for a syllable). Nowadays this type of writing is called *abugida* or *alphasyllabic*, names that imply that in such a writing system the alphabetic and syllabic principles are combined. The invention of the *abugida* type of writing with the akshara as its grapheme was an ingenious achievement of Indians.

The term *grapheme* was proposed by I.A. Baudouin de Courtenay in 1912.¹ Yet nowadays it is used in different meanings. I use the term *grapheme* for designating the independent full-formed unit of linear script which refers to a linguistic unit. As full-formed signs standing in line, graphemes are opposed to diacritics as these are generally placed above or under the line. As independent signs, graphemes can be used separately and have their own linguistic referents; thus, they are opposed to sub-graphemes: to both diacritics (even when placed in line) and parts of ligatures. As signs referring to linguistic units, graphemes are opposed, firstly, to punctuation marks (signs of division and integration), and, secondly, to diacritics that refer to features and characteristics of linguistic units and form modifications of an invariant sign. Thus, a grapheme is considered as a sign, with its signans and signatum. There exists another approach, developed by A.A. Zaliznyak in his theory and practice of deciphering the Birch Bark Manuscripts from Novgorod. He considers a grapheme to be an external side of a sign, a graphic unit, opposed to other graphic units, and capable of being a minor distinguishing element for a pair of written words (cf. the notion “phoneme”), even a diacritic mark.² He uses the term “letter” (*буква*), the more common term in the Russian tradition, for designating a graphic sign. I think that both approaches are appropriate, but the former is more convenient for typological generalization.

Like any other linguistic unit, a grapheme can have its allographs, such as uppercase letters. (When a grapheme is regarded as an external form, then uppercase and lowercase letters are considered different graphemes, but the same letter, as in *China* vs. *china*.)

Due to its structure, a grapheme may be elementary, complex, or compound, formed of several meaningful graphic elements composing a unit.

¹ See BAUDOUIN DE COURTENAY 1912. The term “phoneme” was also developed by Baudouin de Courtenay and his student N.B. Krushevsky.

² Cf. ZALIZNYAK 2002; USPENSKY 1964.

Graphemes, their elements, sub-graphemes, and punctuation marks can be generally called “writing symbols”, a name that indicates the sphere of their use. (There are other categories of writing symbols in morphosyllabic systems that we won't consider here.)

In different types of phonological writing systems a grapheme can have a different linguistic referent – syllabic or alphabetic. Five classes can be noted:

1. Elementary graphemes represent different syllables and even more complex fragments as inseparable units (e.g., CVC, CCVC, CVCC, CVCVC);
2. Elementary graphemes represent only simple moras ((C)V, (V-)C);
3. A 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 (where *x* designates an unspecified vowel);
4. A single grapheme stands for syllables with the same consonant, whereas its modifications differentiate vowels (C^v);
5. Different graphemes represent separate sounds – C and V.

The first four types refer to syllabic writing. The first one is common for systems developed from logosyllabic (morphosyllabic) writing or invented, such as the Yi;³ it can be labeled as the *Yi* type. The second one is the *Kana* type (the Japanese model).⁴ The third one is the *abjad* type (Daniels's term⁵) which can structurally be described as a consonant alphabet, since different graphemes represent different consonants irrespective of vowels; functionally, it has a value of syllabic writing since it allows to read graphemes as syllables (according to Gelb). The fourth is the *abugida* type (Daniels's term) which can structurally be regarded as *alphasyllabic* (Bright's term⁶). The fifth type corresponds to alphabetic writing.

Notionally, it is possible to discern a sixth model with the vowel grapheme being differentiated by a consonant (V^c). The Pahawh Hmong script (invented in 1959)⁷ comes closest to it: Though usually there are two graphemes for a syllable CV in this writing system, their order in the 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 combines the alphabetic and syllable modes of writing: It has full-sized letters for vowel and consonant phonemes, and its syllabic blocks (arranged in non-linear fashion) follow each other in linear order.

The first four possible strategies can be illustrated in the conventional transliteration of the Latin example:

JUVENES DUM SUMUS.

³ The standardization of the logosyllabic Yi writing for the Nosu language in 1980s results in a syllabic system of 819 basic characters (BRADLEY 2009: 179).

⁴ The name moraic for this type (Sproat includes also Linear B, Sumerian, and Mayan writing systems, see ROGERS 2005: 274) does not seem 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, an assumption which remains debatable.

⁵ DANIELS 1996: 3–17, also 2009.

⁶ BRIGHT 1996: 384–391; 2000: 63–71. Cf. COULMAS 1996: 252.

⁷ Cf. ROGERS 2005: 260–263; IVANOV 2004: 25.

The first strategy could have different realizations depending on the set of graphemes, for example, it can be accomplished by 6 separate graphemes:

(you) * (ven) * (es) * (doom) * (soo) * (moos),

or (you) * (ve) * (nes) * (doom) * (soo) * (moos).

The second strategy (the kana type) might have the following representations (irrespective of long or short vowels):

(you) * (ve) * (ne) * (-s) * (du) * (-m) * (su) * (mu) * (-s)

The third one (the abjad type) would be the following, with 9 characters like the previous one, but more difficult to read because of its undetermined vocalism:

J^x * V^x * N^x * S^x * D^x * M^x * S^x * M^x * S^x,

or JVNSDMSMS (C = C^x)

The fourth script (the abugida type), will produce the following sequence, with 9 graphemes as in the abjad type; but the symbols are more complicated, provided with vowel marks:

J^u * V^e * N^c * S^o * D^u * M^o * S^u * M^u * S^o

The fifth type (the alphabetic type) is represented in the Latin itself – with 15 characters (see above).

So, these strategies differ in definitiveness of reading, possible variations in writing, length of the linear chain, complexity of graphemes.

In historical development there is no direct succession. The adaptation of a writing system to a language (the most common way of the development of writing systems), in fact, depends deeply on the structure of the language – its phonological, morphological, and lexical values. Thus, the function of the abjad is appropriate for the phonological, morphological, and lexical structure of classical West Semitic languages, where:

a) A vowel could not open a syllable, so it could be regarded as a dependent characteristic of the syllable;

b) Vowels were variable in word-formation, so they were not a constant characteristic of a root.⁸

Yet, a vowel accompanied a consonant, and, therefore, it was a regular and implied characteristic of a consonant or a syllable as a pronunciation unit as well. So the abjad writing has the advantage of laconism for a scribe, yet it provides difficulties for an inexperienced reader.

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

a) Since a vowel could correspond to a syllable, it needed a special designation;

b) Different vowels needed different designations.

⁸ Reference can be made to LYAVDANSKY 2009.

Thus, systematic vowel differentiation was a main achievement of Indian scripts. Based on akshara graphemes the scripts developed paradigms of akshara's modifications with sub-graphemes that may be named THE AKSHARA'S GRAMMAR.

3. THE ABUGIDA PROPERTIES

The base structure of a writing system is its alphabet.

In order to distinguish between the paradigmatic system and its functioning, I prefer to use the term "alphabet" for a paradigmatic system of graphemes that:

- a) correspond to the sounds – phonemes or syllables of a given language,
- b) are arranged in a stable order,⁹
- c) usually have appropriate names – a fact that makes them LETTERS.¹⁰

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. According to the abugida principle, each syllable should be represented by a single grapheme. The basic syllable forming the alphabet for many abugida scripts is Ca, where C codes a consonant and /-a/ is a constant inherent vowel (in fact, there are other systems with other basic vowels: /-o/ for the Bengali and Thai script, /-ä/ for the Ethiopian script). The main component of an alphabet is a set of Ca graphemes ("consonant aksharas").

Thus, the name of the national Indian hymn "Jana Gana Mana" (that means 'All men's mind') can be written in the Bengali script (where it sounds *jono gono mono*) with a chain of 6 elementary graphemes (only 4 different items), each of them representing a basic syllable:

/jo-no-go-no-mo-no/

জন গন মন

The abugida transliteration may look like [J° N° G° N° M° N°]

There can be two sorts of modification due to the syllable structures of the language:

1. indicating vowel variation (Ci, Cu, Ce, etc), or
2. forming consonant clusters (CCa or CCCa).

They form the enlarged repertoire of graphic signs. The former modification is usually realized as a diacritic modification of an invariant Ca grapheme; the latter as a ligature of two (or more) graphemes. Diacritics usually save the basic shape of a grapheme intact, using the vertical dimension (as superscripts or subscripts); ligatures can enlarge the horizontal dimension of a grapheme, uniting two shapes in a common space in line.

This is the most common model. Yet it has variants and deviations in the historical development.

⁹ 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).

¹⁰ In fact, their names can issue from their pronunciation as in Latin or modern Cyrillic alphabets, yet they become not only marks for sounds, but names of letters (they differ from hieroglyphs which can render sounds or words or morphemes).

Thus, vowel transformations can also be indicated without diacritics, but in changing the orientation of the main graphic shape, as in the Canadian Cree script invented by J. Evans in 1840 (Fig. 1).

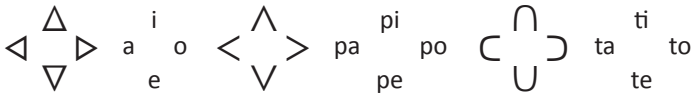


Fig. 1. Vowel transformations in the Canadian Cree script¹¹

In general, vowel modification may be compared with word formation, while consonant ligatures correspond to compounding.

4. THE AKSHARA'S GRAMMAR

The akshara's grammar can be regarded as a model underlying the abugida functioning of Indian scripts.

4.1. VOWEL MODIFICATIONS

The development of the paradigmatic system can be regarded in historical perspective.

Vowels in abugida writing are represented differently in dependent and initial positions. While they are embodied as separate graphemes in the initial position, they are only marked by sub-graphemes in dependent position (after consonant).

In the early scripts, Brahmi and Kharosthi, vowel modifications are marked with a minor graphic element: adjacent or tangential to the main sign WITHIN ITS OWN SPACE in the line. It may change direction or point of contact in accordance with the shape of the grapheme (cf. Ku : Lu in Brahmi, Fig. 2), yet it preserves its distinctiveness. We shall call it a GRAPHON, or inner diacritic. Its function is to change the vowel in a syllable; each graphon indicates similar modifications in different consonant aksharas. This is the first attempt to create the matrix of vowel differentiation, though not quite uniform. A similar system was invented for the Ethiopian writing in the 4th century AD, possibly under the influence of Brahmi (DANIELS 1996; ROGERS 2005: 120). The sequence of the first 4 letters' names produced the term *Abugida* (cf. Alphabet); now the order of letters is different and they form the name Halehame.

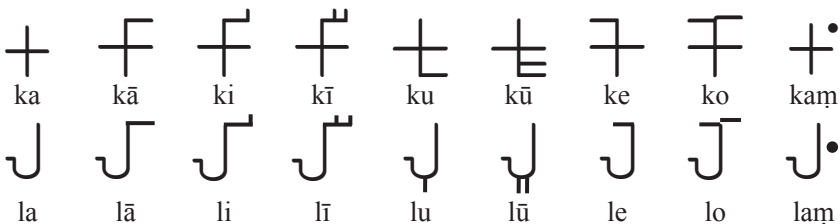


Fig. 2. Paradigm of vowel differentiation for Ka and La in Brahmi¹²

¹¹ Image from http://en.wikipedia.org/wiki/Eastern_Cree_syllabics.

¹² Fragments are extracted from the illustrations at <http://www.omniglot.com/writing/thai.htm>.

A graphon is closely tied to the main sign (radical), thus forming an integral sign. This device, used in early writing systems,¹³ was later developed by separating two parts: a dependent vowel symbol, separate from the main sign, and the main sign itself, which remains unchangeable. Such a way of writing allows the main grapheme to be conserved intact in its own space in line and creates a supplementary space for the sub-graphemes. These symbols differ from graphons and may be regarded as stable alienable marks of their phonetic function — i.e. as pure diacritics, or *matras*. They can create a more uniform matrix of vowel modification, which makes the script more distinctive, as we can see in Devanagari.

Diacritics use the “outside” space of the letter (akshara) – usually above or under the line (superscripts and subscripts); yet in Devanagari there appears also a complex mark for dependent vowels formed by a full-letter height stem (vertical stroke, which is a mark for the long phoneme /a:/ when placed after an akshara) serving as a base for some superscript vowel diacritics. (In some way this mark resembles the Arabic *alif* in its function as 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. So according to its position, the vertical stroke can be a POSTSCRIPT or a PRESCRIPT. Complex diacritics become more common in the South Indian scripts for rendering languages rich in vowels (sometimes forming CIRCUMSCRIPTS and even more complex combinations). Complex diacritics can take even more space than the basic akshara, framing it like “garments” around the “body”, as sometimes in the Tamil (Fig. 4a below) or Thai scripts. The whole complex akshara takes the appearance of an emblem (rather than a linear sign), where the non-linear combination of its elements corresponds to the linear sound sequence in a syllable.

It should be realized that vowel differentiation of a syllable in abugida is not confined to the choice of a mark (inner or outside diacritic) designating a concrete vowel, but represents sequential operations of suppressing the inherent vowel (/a/ from Ca) and then adding another vowel. Thus, a vowel mark is an operator that creates a new unit.

Now, two types of 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 the outside (diacritics). Their common characteristics are their exclusively dependent use and their size, which is usually smaller than that of a main grapheme.¹⁴ 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 phonetic variants of an invariant unit – determines its position (usually outside the inner space of the grapheme) and size (usually not equal to that of the grapheme horizontally and/or vertically); its mode of conjunction with a grapheme is distinct. Full-sized representations of vowels placed in line but differ-

¹³ The device of modification of a sign in its inner space is still used in more recent writing systems based on Latin and 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, as they are found in Vietnamese vowels, as ‘integral parts of the letters’ (DANIELS 2006: 19). Such marks can be named quasi-graphons or quasi-diacritics, if needed.

¹⁴ This is not the case in South and Southeastern scripts with a large number of vowels.

ent from their initial forms (such as, for example, marks for /e/ in Bengali, Tamil, and other languages) 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), Fig. 3.

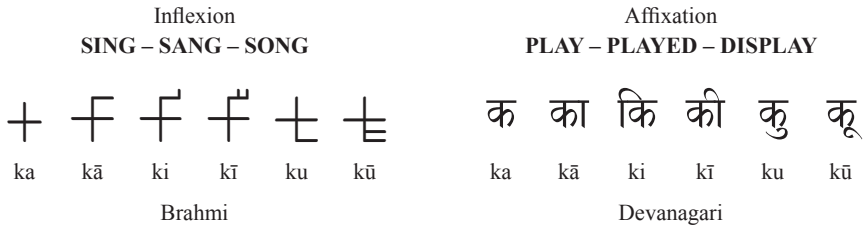


Fig. 3. Comparison of devices in word morphology and akshara-formation

A more complex system of non-initial vowel representation can be found in Modern Tamil. There exist vertical (for short and long *i, u*) and horizontal vowel marks (Fig. 4a). While horizontal diacritics usually have separate shapes (as quasi-graphemes), the former are often united in tight fusion with the main akshara. Such inseparable shapes tend to function like 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 (one standard and three non-standard forms for /u/, (Fig. 4b)), though the same form can refer to different vowels (Fig. 4c). Thus, we can speak about irregular graphons there. All this helps in developing the horizontal chain of characters, though it does not correspond to a linear sequence of sounds (because of the nature of quasi-graphemes). This is the common tendency for South Indian scripts.

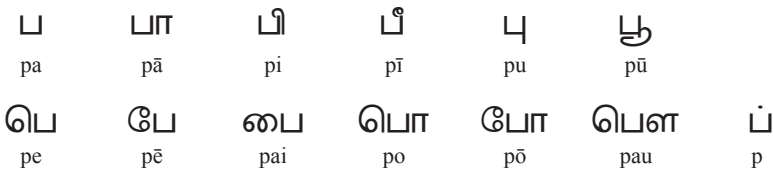


Fig. 4a. Tamil: Regular vowel marks



Fig. 4b. Tamil: Aksharas with 3 non-standard vowel marks



Fig. 4c. Tamil: Alloforms for /-a:/ and /-u/¹⁵

¹⁵ Fragments are extracted from the illustrations in: <http://www.omniglot.com/writing/tamil.htm>.

4.2. INDEPENDENT VOWEL REPRESENTATION

Represented as sub-graphemes or quasi-graphemes in consonant aksharas, vowels have full-sized graphemes for independent positions – in fact, for syllable-initial position. This is one of the basic characteristics of abugida as alphasyllabary. There exist different strategies of initial vowel representation.

4.2.1. THE “EMPTY CONSONANT” DEVICE

This device was exploited (first?) in the Kharosthi script¹⁶ (used mostly for Gandhari, a northwestern Prakrit dialect). There are full-sized graphemes for vowels in the initial position. 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. /ʔ/ for glottal stop, with an inherent vowel /-a/ (as in other invariant graphemes) or an “empty” consonant, so the whole alphabet (*arapacana*) winds up having a uniform structure of graphemes – CV. The invariant form with

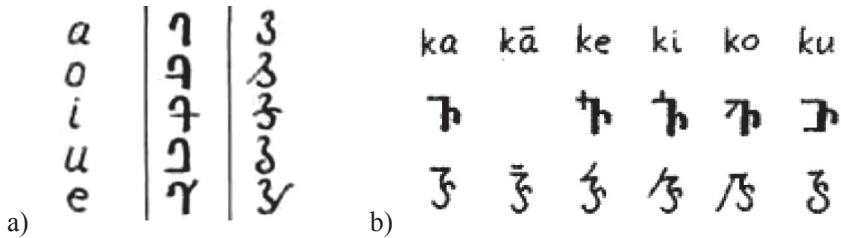


Fig. 5: a) Independent vowels in Kharosthi (two styles); b) paradigm of vowel differentiation for the syllable Ka¹⁷

unpronounced (“empty”) consonant can be modified by graphons in order to represent different vowels, and the same graphons mark parallel modifications of CV aksharas (Fig. 5). Thus, the akshara-grapheme receives the internal structure of an invariant base and a variable graphon; the paradigmatic structure of a system has two components – an alphabet of Ca graphemes¹⁸ (the invariant shapes) and a matrix of their modifications with graphons.

The device of “empty consonant” is also used in some other alphabets; e.g. Tibetan, Thai, and Lepcha use outside diacritics for akshara modifications. An “empty” consonant is there as an unpronounced sign, serving as a carrier for independent vowels and thus maintaining the structural uniformity of graphemes. (As Hans Henrich Hock brings to my attention, a similar attempt was made in Marathi and maybe other modern languages, where the basic symbol for /a/ was modified by the same devices that serve to indicate vowel differences with consonant aksharas.)

¹⁶ The problem of the origin of Indian scripts is reviewed in SALOMON 1995. Kharosthi is generally seen as an invention under the influence of the Aramaic script and probably an antecedent to Brahmi.

¹⁷ Fragments are extracted from the illustrations in: <http://www.rbardalzo.narod.ru/novosti.htm>.

¹⁸ There were some CCV graphemes also (*kṣa, sta, ...*), as elementary shapes and not ligatures.

4.2.2. INDEPENDENT VOWEL SHAPES

Another model of independent vowel representation is exploited in the Brahmi script. There are different graphemes for short vowels, and modifications of these for long vowels (Fig 6).

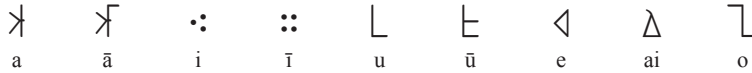


Fig. 6. Independent vowels in Brahmi

In fact, there are only three short vowels: /a/, /i/, /u/, all others are long. The graphemes [u:] and [o] use the common base of the [u]-grapheme with different graphons. So, the following transformations can be seen: [a] => [a:], [i] => [i:], [u] => [u:], then possibly [i] => [e], [u] => [o], and [e] => [ai].

The role of graphons in modifications of vowels (/V/ vs. /V:/) and of consonants (C^{v1} vs. C^{v2}) is entirely comparable. They could be conceived as the operators of (almost) the same order, in contrast to the dot (*anusvara*) which is marking a secondary pronunciation feature (of nasalization). Such a function can be fulfilled by sub-graphemes of the second level. The minimal shape (dot) may show that this feature was perceived as secondary in the phonological hierarchy, and the phonological system could have the following levels: 1) consonant-dominant syllables, 2) vowel-dominant syllables (vowel phonemes), 3) vowels as features of syllables, 4) features of vowels – prolongation, nasalization, final aspiration. Only 1) and 2) correspond to graphemes, while 3) and 4) are represented as operations, marked by sub-graphemes.

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

Finally, the paradigm of vowel differentiation was completed with the introduction of virāma (or halanta) in Devanagari, for designating the “zero” vowel, or the operation of suppression of /-a/. Nowadays, in Hindi, a halanta is often omitted in writing, leaving the true choice up to the competence of the reader.

The Brahmi model was also used in many other Indian scripts: Tamil, Bengali, Oriya, Gurmukhi, etc.

The Kharosthi and the Brahmi models seem to be in opposition, but there are intermediate, compromise devices, so the opposition is not binary.

4.2.3. INTERMEDIATE DEVICES

In the Tibetan script, there are two graphemes for /a/ (há/?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 distinguishing long vowels from short ones (Fig. 7).

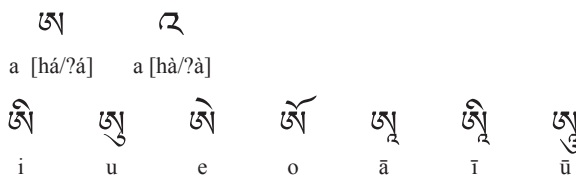


Fig. 7. Some Tibetan independent vowels

In fact, the combination of both signs represents a vertical ligature that serves as base for long vowels. Thus, the Kharosthi model is developed in a more effective structure.

Another intermediate case is the Gurmukhi script. It follows the Brahmi model, in that it has independent vowel graphemes, but they are based on three different graphic shapes (bases for /a/, /i/, /u/) for 10 initial vowels (divided in groups: back, front, and low). Only one of them (the one for /a/ or schwa) can be used independently; the others obtain their different values with the same diacritics as the ones used for vowel designations with consonants, so they can be regarded as vowel-type carriers. Thus, the Gurmukhi script is a compromise and has a more subtle way of vowel representation. T.J. Bhatia regards this mode as a result of generalization, which Devanagari failed to capture (BHATIA 2003: 181–213).

So, the graphic structure of Gurmukhi shows the differentiation of vowels not only for the opposition “short : long”, but also for the finer phonological opposition “front” : “central” : “back”.

4.3. THE DEVELOPMENT OF VOWEL REPRESENTATION

The development of vowel representations in the Brahmi-derived systems shows a tendency to extend the linear chain of graphic elements (especially in South and South-Eastern scripts for languages with a great number of vowels). While the main opposition of high and low vowels /i/ : /u/ is usually represented by super- and subscripts, there is an increasing number of in-line vowel marks. They still remain diacritics since they are dependent on main (consonant) signs and different from the initial vowel forms. Therefore, the vocalization can also here be regarded as a phonetic feature of the syllable (the main phonetic unit), generating its variants – the proper function of diacritics.¹⁹

Thus, the Indian writing systems rely on the syllable, or maybe on the mora, as a main natural pronouncing unit, although not an elementary one, with something like a “head and modifier” structure, where a vowel represents a modifier – “a soul” (in traditional Tamil figurative terms). The set of invariant syllabemes, or the set of “moremes”, is the basis of the abugida system. While such notions as phoneme, morpheme, syllabeme have been elaborated by linguists 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.

Some gradual changes in the representation of vowels can be seen, moving from abjad to alphabet. While in the abjad, vowels are concealed, in so far as any vowel differentiation is absent (if it isn't specifically expressed by means of “matres lectionis”), abugida systems allow consonants to take different shapes “clothed” by diacritics designating different vowels. Ultimately, alphabet systems represent vowels by letters equivalent in size and position to consonant letters. The Greek “claim to democracy” gives vowels their independent status in the Greek alphabetic script and in systems derived from it.

Of course, this is only a metaphor. The 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 Greek phonematic hearing). Abugida scripts are no less exact

¹⁹ This can generate errors among pupils in phonemic analysis of vowels, as is shown in WALI et al. 2009.

in representing sound sequences; and the degree of exactness is relative even in modern alphabets. The absence of definite vocalization in abjad systems is more appropriate for maintaining the recognizability of roots. So the question of progress is more rhetorical than prognostic.

4.4. CONSONANT CLUSTER REPRESENTATION: LIGATURES

Vowel diacritics produce complex aksharas, which can be arranged in paradigms. This resembles grammatical paradigms in morphology, where affixes modify grammatical forms. When speaking about the akshara's grammar, the set of syllabemes may be taken as its "lexicon" and diacritics of vocalization may be taken as grammatical elements since they produce standard modifications.²⁰

Another process is a compound akshara formation: It is to a large extent a matter of the "lexicon" and not only of "grammar". The integration of two (or more) aksharas helps to represent consonant clusters, and as a result the akshara becomes compound, with its components having more or less equal status. Compounding is a way to avoid uncertain readings of akshara sequences: Before the virāma (cancelling the inherent vowel) had been introduced, the akshara allowed sometimes a reduced, consonantal reading. Ligatures were the constructive technical solution of the problem. Thus, a LIGATURE is regarded here as an integrated shape of two (or more) independent graphemes of equal status in a common space (without a lacuna).

The Brahmi script has a model for constructing ligatures, or conjunct consonants (*samyuktākṣara*), to represent consonant clusters (which were rare in Pali²¹). The model of conjunction is vertical, the base could be chosen as the best graphic variant. A ligature does not differ greatly from the akshara with a vowel diacritic: It forms an integral sign in a line.

The technique of combining 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; they have one common vertical stroke (if any), which is not longer than the line; that is why the vertical ligature is usually compressed. Horizontal ligatures are not compressed, but only have a common stroke. 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, the consonant ligature differs from the vowel modified akshara as a unique integrated sign within the line. The whole conjunct tends toward minimal extension of its space and exploits not the principle of linearity (connection in succession), but the emblematic principle of a balanced graphic composition. Vertical outline space is left for vowel designation in aksharas, while horizontal extension is common for ligatures and some vowel elements.

In Tibetan, the rules for conjunction are more complicated; sound sequences are not so balanced there and the structure of a syllable can be (CC)CV(CC) with a maximum of three consonants before the vowel and two after it. As a consequence, the Tibetan script has a two-dimensional way of representing consonant clusters within a syllable: (1) vertical stacks of

²⁰ In reconstructing the akshara's grammar I follow the semiotic approach, formulated by M. Neeff: '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).

²¹ See SHEVOROSHKIN 2004: 135–138.

graphemes can be completed with one to two graphemes in line; (2) conjunct elements may be placed before and after, above and/or below the central one, which is the “root” of the 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). Vowel diacritics are placed above or below the root. The Tibetan script does not use compression; so horizontal elements of a syllable form a free sequence, and the problem of dividing closed syllables arises. The problem is solved by introducing a dividing dot after 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 value, 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, while the remaining signs in line can be regarded as graphemes (aksharas) with reduced phonetic meaning in a complex-compound akshara — an orthographic syllable. The economy of effort allows omitting the dots; resolving the resulting uncertainty is left to the competence of the reader. Nowadays, pronunciation differs greatly from orthography (many consonants are unpronounced, see Fig. 8), but this is not a graphic complexity, but an orthographic one.

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

Fig. 8. Some Tibetan numerals with transcription and transliteration²²

In many Indian scripts the vertical dimension serves for conjunct consonants (often in compressed form), while many vowel diacritics are placed in line; as a result, the paradigmatic system can have an additional component of conjunct consonants. Another way to represent consonant clusters is to use a *virāma* (Tamil *pulli*) in the non-final akshara.

5. LINEAR AND EMBLEMATIC MODES OF WRITING:
COMPARISON BETWEEN PHAGS-PA
AND THE ALPHABETIC KOREAN WRITING

The extension of the vertical dimension used in Tibetan results in the uniquely vertical direction of writing in Phags-pa script, which is derived from Tibetan but culturally affiliated to Chinese script in its appearance (Fig. 9). The Phags-pa script deviates completely from two-dimensional writing. 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 reserved in two main points:

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

²² Transliterations given according to: http://tibetan.bitecs.ru/files/text_intros.pdf.

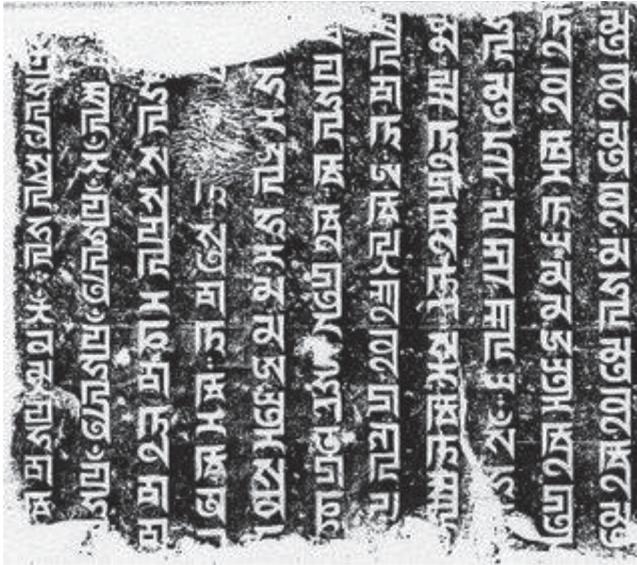


Fig. 9. Phags-pa script

In the syllable block the problem of finding a “root” component is more serious than in Tibetan, since the model of conjunction can be fusion rather than distinct agglutination.

Another way of consonant-vowel arrangement in a syllable block was elaborated for Hangul, the alphabetic Korean script (1443–1444). The problem of “root” does not exist here since all letters have equal alphabetic value. The most widespread structures of graphic syllables are CV or CVC (they do not always correspond to the phonetic syllables of modern Korean). So the syllable block includes two or three components, but the relations between them are coordinate rather than subordinate as in aksharas, since their components are letters. A consonant may seem more important since it always opens a syllable even if it is absent in the pronunciation. Yet, a vowel has a somewhat privileged graphic role, as its outline (vertical or horizontal) defines the mode of conjunction of a consonant letter (or letters), which can be compressed (in the horizontal or vertical dimension respectively). The vowel-consonant arrangement in a block is determined not by linear order, but by the two-dimensional emblematic principle of a balanced graphic composition and can be compared with the akshara arrangement. Yet the Korean script is more transparent than the Pags-pa script.

Compositionality is the most important property of Korean writing: Though looking like integrated signs, its blocks can be separated into letters, and letters into graphic elements that may often have phonetic meaning (distinctive features). Another property is the symbolic and iconic character of their outlines that encode their articulation. These qualities allow defining the Korean script as “featural” writing (DANIELS 1996; LEE 2009), where generalization is more perfect than in any other existing system. I would say that the Korean script combines an emblematic mode of syllable presentation with alphabetic elements and a linear order of syllable blocks. Yet the Korean script does not represent the exact pronunciation, because of some graphic and orthographic conventions. Some of these are connected

with syllable structure: A syllable beginning with V receives CV representation, and CV-CV may be represented as CVC-CV; see the following examples (Tab. 1).²³

Table 1. Comparison of syllabic and graphic structure in Hangul

Hangul	Syllables (transcription)	Syllabograms (transliteration)	Translation
우수	u-su	ngu-su	rain's water
곡우	ko-ku	kok-ngu	crops rain
립하	ip-xa	lip-xa	beginning of summer
립동	ip-tong	lip-tong	beginning of winter

Since a vowel cannot stand in initial position, an empty consonant is used (which is pronounced as /-ng/ in the final position); the grapheme for /l-/ is not pronounced in initial position, but the letter refers to the traditional Chinese syllable /li/. Graphic syllables (syllabograms) may correspond to the morphological structure of a word rather than to its pronunciation structure, as it is for <kok-ngu>. Besides, the pronunciation of some consonants depends on their position in the word (they have allophones). Historical changes also have taken place. So, different rules interact in the writing, which makes it not as easy as it seems.

Moreover, Hangul is not the only script in use in South Korea. In fact, the writing system is of a mixed nature, including Chinese characters in a linear row (horizontal from left to right, or vertical, in stacks from right to left). The use of Chinese characters seems to be not only a case of tradition or ideology, but also serves for easier encoding of lexical notions.

So the Korean script combines the linear principle in the sequences of blocks (syllabic or hieroglyphic) and the emblematic principle²⁴ of arrangement of letters in the syllable block. It provides a simpler solution to the representation of consonant clusters than the abugida system of Phags-pa. Yet the latter was conceived as a graphic system for quite different languages (Mongolian, Tibetan, Chinese, Uyghur, Sanskrit) and satisfied its tasks. Nowadays a globalized version of Hangul, Nurigeul, has been proposed for endangered languages of Southeast Asia as an international graphic system.²⁵

6. ANALYTIC VS SYNTHETIC WRITING

The abugida system differs from other types in the use of two-dimensional writing. Its grapheme represents a hierarchical composition encoding the hierarchy of phonological units and features.

Simple graphemes – consisting of one “radical” for Ca – can be distinguished from complex and compound ones, combined with diacritics, or bound in ligatures. It is possible to draw a parallel between this difference and the analytic/synthetic character of word sequences (in accordance with the typological approach to language analysis by Greenberg²⁶):

²³ These are the traditional names of calendar seasons.

²⁴ See more on the emblematic principle of writing in FEDOROVA 2009.

²⁵ See KIM-CHO 2008.

²⁶ GREENBERG 1960: 178–194; 1963: 60–94.

Analytic chains do not integrate lexical and grammatical elements, 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), while synthetic writing presupposes the use of complex and compound graphemes. This opposition allows speaking about the degree of synthesis (complexity) of a grapheme (S_g) and the degree of synthesis (graphic complexity) of a writing system (S_w), which can be calculated.²⁷

Alphabetic writing is subordinate to linear order, it can be qualified as analytic. The abugida scripts for different languages can differ in the degree of graphic complexity. The Korean writing should be appreciated in its linear units, the syllable blocks; in the degree of synthesis it can be compared with abugida writing.

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 (or syllable) division marks. Finally, writing styles (cursive, calligraphic, or printed) can differ greatly in this respect. In addition, graphic complexity does not correspond to orthographic complexity which is connected mostly with language change.

7. CONCLUSION

This survey, as superficial as it is,²⁸ yet reveals the tendency that exist in the Brahmi-derived scripts to use two-dimensional structures in the representation of sound sequences. Though a linear order of analytic script seems to be an easy solution for exact representation, compact forms of complex signs have their advantages. They can refer to different structures of a syllable and represent them adequately using two dimensions: horizontal and vertical. The meaningful elements of writing are arranged according to their value: syllable, sound as a part of a syllable, phonological feature of a sound. The paradigm of akshara modifications forms a grammar that is realized in abugida writing. Even in the seemingly linear graphic chains there are quasi-graphemes or inverse combinations that make the script more complex; vertical representation in Phags-Pa also combines linear order (of stacks) and block integration. So the emblematic nature of akshara embodies the possibilities and advantages of abugida writing.

We can see a parallel tendency of two-dimensional arrangement in the alphabetic writing of Southeast Asia. For example the Korean Hangul and Pahawh Hmong systems prefer to combine letters in syllable blocks with vowel letters as basic elements in non-initial position. Such representations break the linear order and create emblematic signs.

Thus, successive phonemic representation in alphabetic writing is not the only way of graphic arrangement. The principle of linearity can coexist with the emblematic principle that reveals the hierarchy in phonological structures. These structures require decoding in reading.

²⁷ An example of a calculation is proposed in FEDOROVA 2012. It shows (in preliminary estimation) the higher degree of synthesis for Tibetan in comparison with Bengali, Devanagari, and Tamil.

²⁸ For a more detailed analysis see FEDOROVA 2012.

The graphic shapes of the akshara can be represented as forming “a grammar” parallel to morphological structures. This gives perspective for estimating its properties in the frame of the akshara’s grammar – or the graphic grammar.

REFERENCES

- BAUDOIN DE COURTENAY Ivan A. 1963 (first published in 1912, in Russian). “On the Relation of Russian Writing to the Russian Language.” In: *Baudouin de Courtenay, Ivan A. Selected Works on General Linguistics 2*: 209–235. Moscow: Academy of Science of the USSR.
- BELOVA A., KOGAN L., LOESOV S., ROMANOVA O. (eds.). 2009. *Languages of the World: The Semitic Languages: Akkadian, Northwest Semitic*. Moscow: Academia.
- BHASKARARAO Peri (ed.). 2003. *International Symposium on Indic Scripts: Past and Future*. Tokyo: Tokyo University of Foreign Studies.
- BHATIA Tej K. 2003. “The Gurmukhi Script and Other Writing Systems of Punjab. History, Structure and Identity.” In: BHASKARARAO 2003: 181–213.
- 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: DANIELS & BRIGHT 1996: 384–391.
- BRIGHT William. 2000. “A Matter of Typology: Alphasyllabaries and Abugidas.” *Studies in the Linguistic Sciences* 30(1), 63–71.
- COULMAS Florian. 1996. *The Blackwell Encyclopedia of Writing Systems*. Oxford: Blackwell.
- DANIELS Peter T. 1996. “The Study of Writing Systems.” In: DANIELS & BRIGHT 1996: 3–17.
- DANIELS Peter T. 2006. “On Beyond Alphabets.” *Written Language and Literacy* 9(1), 7–24.
- DANIELS Peter T. 2009. “Two Notes on Terminology.” *Written Language and Literacy* 12(2), 277–281.
- DANIELS Peter, BRIGHT William (eds.). 1996. *The World's Writing Systems*. New York: Oxford University Press.
- FEDOROVA Liudmila L. 2009. “The Emblematic Script of Aztec Codices as a Particular Semiotic Type of Writing System.” *Written Language and Literacy* 12(2), 258–276.
- FEDOROVA Liudmila L. 2012. “The Development of Structural Characteristics of Brahmi Script in Derivative Writing Systems.” *Written Language and Literacy* 15(1), 1–25.
- GREENBERG Joseph H. 1960. “A Quantitative Approach to the Morphological Typology of Language.” *International Journal of American Linguistics* 26(3), 178–194 (Russian version: ГРИНБЕРГ Дж. 1963. “Квантитативный подход к морфологической типологии языков.” Звегинцев В.А. (ред.). *Новое в лингвистике* 3, 60–94. Москва, изд. Иностранная литература.).
- IVANOV Vyacheslav V. 2004. *The Linguistics of the Third Millennium: Questions to the Future*. Moscow: Languages of Slavic Culture (in Russian).
- KIM-CHO Sek Yen. 2008. “A Solution to Save Endangered Language.” In: *The 18th International Congress of Linguists. July 21-26, 2008, Korea University, Seoul, Korea. Abstracts*, vol. 2: 178–179.
- LEE Sang Oak. 2009. “The Korean Alphabet: An Optimal Featural System With Graphical Ingenuity.” *Written Language and Literacy* 12(2), 202–212.
- LEE Sang Oak (ed.). 2010. *Contemporary Korean Linguistics: International Perspectives*. Gyoha-eup (Korea): Thaeaksa.
- LYAVDANSKY Alexej K. 2009. “The Origin and Early Development of the West Semitic Alphabets.” In: BELOVA et al. 2009: 811–821.
- NEEF Martin. 2010. “On Defining Phonology.” In: LEE 2010: 216–230.
- PANDEY Pramod. 2003. “Phonetic and Phonological Bases of Hindi Orthography.” In: BHASKARARAO 2003: 41–61.
- ROGERS Henry. 2005. *Writing Systems: A Linguistic Approach*. Oxford: Blackwell.
- SALOMON Richard G. 1995. “On the Origin of the Early Indian Scripts: A Review.” *Journal of the American Oriental Society* 115(2), 271–279.
- SALOMON Richard G. 1996. “South Asian Writing Systems: Brahmi and Kharoshthi.” In: DANIELS & BRIGHT 1996: 373–383.
- SHEVOROSHKIN Vitaly V. 2004. *The Sound Chains in the Languages of the World*. Moscow: URSS (in Russian; 1st ed. 1964).

- SPROAT Richard. 2006. "Brahmi-derived Scripts, Script Layout, and Segmental Awareness. *Written Language and Literacy* 9(1), 45–66.
- USPENSKY Vladimir A. 1964 – В.А. Успенский. "Одна модель для понятия фонемы." *Вопросы языкознания*, 1964: 6, 39–53 (in Russian).
- WALI A., SPROAT Richard, PADAKANNAYA P., BHUVANESHVARI B. 2009. "Model for Phonemic Awareness in Readers of Indian Script." *Written Language and Literacy* 12(2), 161–169.
- ZALIZNYAK Andrej A. 2002. – А.А. Зализняк, "О понятии графемы" // А.А. Зализняк. *Русское именное словоизменение. С приложением избранных работ по современному русскому языку и общему языкознанию*. М.: Языки славянской культуры, 2002: 559–576 (in Russian)

WEB MATERIALS

- <http://www.rbardalzo.narod.ru/novosti.htm>.
<http://www.omniglot.com/writing/tamil.htm>.
<http://www.omniglot.com/writing/thai.htm>.
http://tibetan.bitecs.ru/files/text_intros.pdf.