Initial literacy in Devanagari: What Matters to Learners

Renu Gupta
University of Aizu

1. Introduction

Heritage language instruction is on the rise at universities in the US, presenting a new set of challenges for language pedagogy. Since heritage language learners have some experience of the target language from their home environment, language instruction cannot be modeled on foreign language instruction (Kondo-Brown, 2003). At the same time, there may be significant gaps in the learner’s competence; in describing heritage learners of South Asian languages, Moag (1996) notes that their acquisition of the native language has often atrophied at an early stage.

For learners of South Asian languages, the writing system presents an additional hurdle. Moag (1996) states that heritage learners have more problems than their American counterparts in learning the script; the heritage learner “typically takes much more time to master the script, and persists in having problems with both reading and writing far longer than his or her American counterpart” (page 170). Moag attributes this to the different purposes for which learners study the language, arguing that non-native speakers rapidly learn the script since they study the language for professional purposes, whereas heritage learners study it for personal reasons. Although many heritage languages, such as Japanese, Mandarin, and the South Asian languages, use writing systems that differ from English, the difficulties of learning a second writing system have received little attention. At the elementary school level, Perez (2004) has summarized differences in writing systems and rhetorical structures, while Sassoon (1995) has documented the problems of school children learning English as a second script; in addition, Cook and Bassetti (2005) offer research studies on the acquisition of some writing systems.

For heritage learners, their introduction to an Indic writing system\(^1\) may take place either in the home or at the university. If the writing system is learned in the home, parents teach them much as they were taught, using textbooks published in India. At the university, the chances are that the traditional approach will be used to teach the script. In the traditional approach, the symbols are presented in the varnamala (phonetic) sequence, followed by the CV
(consonant-verb) combinations, and then the CC (consonant-consonant) ligatures before learners proceed to the serious business of reading. This is the method employed in school textbooks for children in India. It is also used in most textbooks designed for foreign language instruction, such as Snell (2003) from the UK or Vikal and Dhingra (2005) from Malaysia. Samples of this approach can be viewed on the websites of several universities that teach Hindi—at Syracuse University in the US, the School of Oriental and African Studies (SOAS) in the UK, and Australian National University and La Trobe University in Australia. An interesting aspect of the material designed for learners outside India is that, in addition to presenting the symbols in the traditional sequence, the pedagogical approach is linguistic; the explanations tap learner knowledge of terms such as velar consonants and inherent vowel. What is missing is a psycholinguistic perspective that addresses learner difficulties in recognizing, discriminating, and remembering the symbols.

If a script is viewed from the learner’s perspective, it can be taught (and learned) easily. One example from foreign language instruction is the book by Mace (1961) that is used to teach the Persian script. Instead of following the traditional sequence of letters—alef, be, pe, te, ...—small sets of letters are taught together. In the first lesson, only three letters are taught—$a$, $n$, and $b$—with which the learner can immediately read and write six words—$ba$ (with), $ab$ (water), $baba$ (father), $an$ (that), $nan$ (bread), and $banna$ (builder)—as well as short phrases ($with$ that $water$). Note that the letters for /b/ and /n/ are visually similar in certain combinations and only the position of the dot differentiates them.

Similar thinking has prompted educationists in India to re-examine how the Indic scripts are taught. When literacy was confined to a small elite group in India, the effectiveness of the pedagogy was never questioned, and in the formal educational system designed for middle-class populations both within and outside India the Indic scripts are still taught through the traditional approach. But faced with the challenges of low literacy rates and high attrition among first-generation learners, educationists have begun to devise teaching material that is more psychologically effective.

From a review of studies, Padakannaya and Mohanty (2004) suggest that reading acquisition in an Indic script follows a “simple to complex” order that is linked to the levels of complexity inherent in the script. The normal sequence is:

Simple basic letters -> letters with vowel diacritics -> letters with ligatures -> complex
Most research studies focus on Stages 2 to 4 (see the following section), i.e., after the basic symbols have been learned. However, learning the basic symbols can present challenges for learners (for the Japanese case, see Tsukada, 2007). This paper is concerned with Stage 1 in which the basic symbols are taught.

The paper first summarizes research on processing issues in an Indic writing system and then examines two characteristics of Indic scripts—the systematic arrangement of the sounds and the presence of an inherent vowel; Sections 4 and 5 describe two non-traditional and relatively recent approaches used to teach the Devanagari symbols. Although these two approaches are not widely used, they offer viable solutions for teaching heritage learners the scripts used in writing several of the South Asian languages.

2. The Psycholinguistics of Indic scripts

Considerable research has been conducted on the difficulties children encounter in learning an alphabetic script, and English in particular. The most critical factor appears to be phonemic awareness (Goswami and Bryant, 1990; Stanovich, 1986), which involves an awareness of sound units smaller than the syllable. This is necessitated by the structure of alphabetic writing systems that map a phoneme to a letter.

Non-alphabetic writing systems, such as the Indic scripts, have received less attention, but recent research indicates that factors implicated in alphabetic reading may not apply to reading Indic scripts (Karanth, 2002; Karanth, Mathew, and Kurien, 2004; Padakannaya and Mohanty, 2004). In Indic scripts, the central unit is not the phoneme but the *akshara* which consists of CV sequences (Patel, 2004); processing is partly syllabic and partly phonemic (Vaid and Gupta, 2002). In a study of children studying in Oriya-medium schools, Fukuzawa & Prakash (1993) found that they went through two distinct stages: decoding and decoding with comprehension. The first stage, or decoding, lasts for several months indicating that it takes time to master the complexities of the script. Research studies that investigated children, adults, and dyslexics reading in different Indian languages and scripts found that specific characteristics of these scripts caused processing difficulties. For example, the vowel allographs or diacritics require additional processing (Anand, 1990; Nehru, 2001; Patel, 2004; Purushothama, 1990) and that this effect is accentuated when the CV combination is written as VC, as in the word कि that is pronounced /ki/ but written <ik> (Vaid, 2002). Other features of
Indic scripts, such as the consonant ligatures, are also likely candidates for high processing loads.

The difficulties in learning an Indic script can be seen in a pilot study that I carried out in India. The data were collected in three schools in a small town in north India, where school children first learn the English script and then Devanagari. By the age of 3 ½, children have learned to write lower-case and uppercase letters, as well as the cursive script in English. Then they move to the Devanagari script; this is taught through the traditional method, namely, the varnamala, followed by the barakhadi chart to automatize the CV combinations and, finally, the most useful conjuncts. According to their teachers, children found Devanagari more difficult than learning the English letters.

Although the strokes in Devanagari differ from the English strokes (Lambert, 1953; see Figure 1), the teachers tapped the children’s knowledge of English to teach them the Devanagari symbols. For example, to learn the shape for अ, children are told that it is like the capital letter B; to learn इ, they are asked to write the letter S and add a tail. This strategy may not be very effective; their copy books showed that instead of writing ऊ, several children wrote it with the tail moving out as ऋ, using this form for several months. The confusion arises from the cursive script in English, where letters have to be joined, as in the letter <h>.

![Figure 1. Strokes used in Devanagari (from Lambert, 1953: page 5)](image)

In addition to interference effects from their first script, symbol formation was clearly a problem. Videotapes showed that the children were forming the shapes incorrectly, by moving the pencil upward to the horizontal line instead of downward from it. Consequently, they had little control over the symbol shapes, making it difficult to distinguish their formation of the symbols थ, य, भ, म, and घ (namely, the symbols for th, y, bh, m, and gh).

In Stage 2, the CV combinations are taught. In line with the research findings cited above, teachers reported that the children found the vowel diacritics especially difficult, the most problematic being ए which is the diacritic for /i/, thereby supporting Vaid’s findings (2002) with adult readers of Devanagari.

Taken together, the research evidence indicates that processing an Indic script is not easy for
either children or adults, and its complexity makes it difficult to learn, regardless of whether it is the learner’s first or second script.

3. Characteristics of Indic scripts

Indic scripts are primarily found in South and Southeast Asia. In India, nine Indic scripts are officially recognized: Bengali, Devanagari, Gujarati, Gurmukhi, Kannada, Malayalam, Oriya, Tamil, and Telugu. Since these scripts are derived from a common source script—Brahmi—their organizing principle remains the same although their symbols differ. Two properties of these scripts concern us here: (a) the sounds are organized systematically based on phonological principles, and (b) each consonant has an inherent vowel.

3.1 Organization of the symbols

In Indic scripts, the symbols are systematically organized based on articulatory phonetics. Vowels and consonants are separated; the vowels are arranged by vowel length (short followed by long), and the consonants are arranged by place and manner of articulation; for example, /k/ and /g/ are placed on the same row because they differ only in voicing.

Table 1 shows the arrangement of sounds for an Indic script along with the symbols used for Devanagari. This arrangement, known as the *varnamala* or *aksharmala*, is read along the rows and is the traditional sequence in which school children are taught the script; they begin with the vowels and move through the chart, learning the sounds and symbols one-by-one. Knowledge of the sequence is necessary because, like the alphabet, it is used to organize items in dictionaries and lists; a more recent application is text messaging on mobile phones, where all the symbols in a row are assigned to one key and multitap is used to access the required symbol (Gupta and Sornlertlamvanich, 2007).
| Vowels | 
|---|---|---|---|---|---|---|---|
| Primary Vowels | a | ā | i | ī | u | ū | r |
| Secondary Vowels | e | ai | o | au | a` | ah | | |

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Voiceless Plosives</th>
<th>Voiced Plosives</th>
<th>Nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unaspirated</td>
<td>Aspirated</td>
<td>Unaspirated</td>
</tr>
<tr>
<td>Velar</td>
<td>kə</td>
<td>kʰə</td>
<td>gə</td>
</tr>
<tr>
<td>Palatal</td>
<td>tʃə</td>
<td>tʃʰə</td>
<td>də</td>
</tr>
<tr>
<td>Retroflex</td>
<td>tə</td>
<td>tʰə</td>
<td>də</td>
</tr>
<tr>
<td>Dental</td>
<td>tə</td>
<td>tʰə</td>
<td>də</td>
</tr>
<tr>
<td>Labial</td>
<td>pə</td>
<td>pʰə</td>
<td>bə</td>
</tr>
<tr>
<td>Semi-Vowels</td>
<td>jə</td>
<td>rə</td>
<td>lə</td>
</tr>
<tr>
<td>Sibilants</td>
<td>shə</td>
<td>sə</td>
<td>sə</td>
</tr>
<tr>
<td>Glottal</td>
<td>hə</td>
<td>ɭ</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1. Phonological Inventory of an Indic script. Each cell shows the pronunciation followed by the Devanagari symbol.*

The *varnamala* makes perfect sense to a linguist and anyone trained in phonetics can
articulate the sounds in each cell based solely on the phonetic description. However, since reading and writing do not involve sounds but symbols, it is easy to recite the sequence but learning the symbols is another matter. To learners, such as children, adults learning a new script, and illiterate adults, many symbols look similar and it is not easy to discriminate one symbol from another. For example, in Devanagari the symbols for the related sounds /b/ and /bh/ are visually distinct (namely, ब and भ), whereas the symbols for the unrelated sounds /b/ and /v/ are visually close (namely, ब and व). Further, the symbol घ looks merely like a stylistic variation of घ, but these are two different symbols that represent different sounds. Such differences can confuse beginners.

Recognizing the importance of pattern recognition, recent methods designed for adult learners in India and the US use a shape similarity approach. Section 4 describes this method as it is implemented by the University of Pennsylvania in the US for foreign language and heritage learners, and by the Central Institute of Indian Languages (CIIL) in India for literate adult learners.

3.2 The Inherent Vowel

Unlike the alphabetic writing system, the sounds in the varnamala are not phonemes because each consonant contains an inherent vowel, such as a schwa. In English, the letter <k> represents the sound /k/, but in Devanagari the symbol क represents the sound /ka/. Hence, a symbol does not represent a phoneme but a CV combination; in order to obtain a phoneme, a virama diacritic is used to ‘kill’ the vowel. For example, the virama is placed below the symbol क to obtain क.

From the perspective of an alphabetic writing system, the inherent vowel in Indic scripts indicates that readers lack phonemic awareness and phoneme segmentation skills. However, phonemic awareness may not be critical in the case of the Indic scripts (Karanth, 2002; Padakannaya and Mohanty, 2004). In fact, the presence of the inherent vowel offers an advantage in teaching a script. Due to this property, it is possible to write several CVC and CVCVC words in Indic scripts by combining consonant symbols, without using additional vowels. For example, learners can create words by combining the symbols for k and l to form kal (tomorrow), or k, l, and m to create the word kalam (pen).

This property is utilized in the teaching material described in Section 5. At Eklavya in India,
instead of learning the *varnamala* sequence, children begin by recognizing and creating words from a sub-set of frequently used symbols. This method need not be confined to the school population, but as can be seen from Mace (1961), it is also an effective teaching approach for literate adults.

4. Teaching through Shape Similarity

Unlike children learning their first script, heritage learners are aware of the features of writing, particularly the notion of sound-symbol correspondence. However, the problem is recognizing and memorizing a new set of symbols. Although reciting the *varnamala* sequence helps learners practice the sounds, it does not help them memorize the symbols. The shape similarity approach attempts to address this issue. Initially devised at the Central Institute of Indian Languages (CIIL) in Mysore, a version for heritage learners is available from South Asia Regional Studies at the University of Pennsylvania.

In the 1970s, the Central Institute of Indian Languages (CIIL) in Mysore began to experiment with alternative methods of teaching the Indic scripts to literate adults (Pattanayak, 1990; Sambasiva Rao, 1978). Instead of the traditional sequence based on sound similarity, clusters of symbols were taught based on their visual similarity; this move was motivated by Lambert’s (1953) work that highlighted the importance of pattern recognition. In her book, Lambert (1953: 7) places similar shapes on the same line:

\[
\begin{aligned}
\text{व, ब, क} & \text{(namely, w, b, k)} \\
\text{ग, न, म, भ} & \text{(namely, g, n, m, bh)} \\
\text{र, स, ख} & \text{(namely, r, s, kh)} \\
\text{ह, इ} & \text{(namely, h, i)}
\end{aligned}
\]

She also draws attention to the shape similarity in the following symbols:

\[
\begin{aligned}
\text{ट, ठ, ड, ढ, द, इ} & \text{(namely, t, tʰ, d, dʰ, d, i)}
\end{aligned}
\]

In the CIIL film, *Learn Devanagari* (1975), all the basic Devanagari symbols (vowels, consonants, and semi-vowels) are grouped into 11 categories based on the similarity of their shapes. In each group, the first symbol is the simplest and requires the smallest number of strokes. Each group of symbols is presented in turn, and the learner’s attention is drawn to visual details that distinguish one symbol from another. For example, viewers are instructed to
“notice the knot in म” and “Notice the line inside the letter ब”.

The version developed at the University of Pennsylvania differs from the CIIL version in that it separates the consonants from the vowels. Table 2 shows the consonant groups for the version from the University of Pennsylvania.

<table>
<thead>
<tr>
<th>Group</th>
<th>Devanagari Symbols</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>व, क, ब</td>
<td>w, k, b</td>
</tr>
<tr>
<td>2</td>
<td>र, ख, स</td>
<td>r, kh, s</td>
</tr>
<tr>
<td>3</td>
<td>त, ल</td>
<td>t, l</td>
</tr>
<tr>
<td>4</td>
<td>छ, ज, झ, न</td>
<td>tf, dʒ, ɲe, ɲə</td>
</tr>
<tr>
<td>5</td>
<td>ग, म, भ</td>
<td>g, m, bh</td>
</tr>
<tr>
<td>6</td>
<td>ढ, प, ष, फ</td>
<td>n, p, s, ph</td>
</tr>
<tr>
<td>7</td>
<td>घ, ध, छ</td>
<td>gh, dh, tʃh</td>
</tr>
<tr>
<td>8</td>
<td>श</td>
<td>sh</td>
</tr>
<tr>
<td>9</td>
<td>य, थ</td>
<td>y, th</td>
</tr>
<tr>
<td>10</td>
<td>ड, ढ, ण, त</td>
<td>ɖ, ɳ, tʃ, dʒh</td>
</tr>
<tr>
<td>11</td>
<td>ट, ठ, ड, ढ</td>
<td>t, tʰ, d, tʰ, ɖ</td>
</tr>
<tr>
<td>12</td>
<td>ह</td>
<td>h</td>
</tr>
</tbody>
</table>

*Table 2. Symbol groups based on shape similarity*

The effectiveness of the shape similarity method is not clear because the results have not been documented, either within India or outside the country. However, there may be some advantages in using this approach. For literate adults, such as heritage learners, many symbols in the new writing system look alike; when similar shapes are clustered, learners can attend to the differences among them. In addition, writing becomes easier because symbols in each group can be written with the same basic strokes but small modifications. To literate
adults who already know an Indic script the sound sequence is familiar, but different the symbols are used; instead of spending time re-teaching the sound sequence, this approach directs attention to the unfamiliar components\(^3\).

With illiterate adults, such an approach has been less successful. The National Literacy Mission (NLM) used this method in the 1980s with illiterate adults, but abandoned it after a few years. Adult learners preferred the traditional sequence, stating two reasons: (a) their friends had told them that this was not the correct way to learn the symbols, and (b) it was easier to learn the symbols in a sequence, because reciting the \textit{varnamala} helped them memorize the symbols (Jennifer Bayer, personal communication).

CIIL continues to use the shape similarity method in its textbooks and distance learning material. Their printed material includes the \textit{Apni Boli} series (1975) to teach several Indian languages. In their recent textbooks, the Indic scripts are taught using two approaches: first, the traditional sequence is taught and then symbols that are often confused are shown in a separate chart (see the Kannada textbook by Kalegowda and Thumbaramatti, 2001). Online courses for four languages—Bangla, Kannada, Tamil, and Telugu—also deploy shape similarity (Central Institute of Indian Languages, 2007). Of course, this results in a different symbol set for each script; the first symbols taught in Bangla are \textit{ta}, \textit{a}, and \textit{a}, whereas in Tamil they are \textit{ta}, \textit{pa}, \textit{ma}, and \textit{ya}.

Although there is no empirical evidence for the effectiveness of this method, it may be more learner-friendly than the traditional approach. Certainly, it opens up a pedagogical and research area in which designers can experiment with different symbol clusters that are tested for their effectiveness with different learner populations.

5. Teaching Clusters Based on Frequency

The inherent vowel in Indic consonants can be used to teach the symbols and this approach is seen in the teaching material devised by Eklavya, a non-governmental organization (NGO) in Bhopal, India. The teaching material was originally devised to meet the needs of the learner population but it is made possible because of this inherent property.

The learners are children who come from the lower socioeconomic strata of society, speak the local dialect rather than Standard Hindi, and are first-generation learners; school attrition rates are high. Reasoning that the children needed to immediately see the relevance of what they
were taught, the organization focused on bringing meaning into the classroom. Since in the
traditional method word reading is delayed until all the symbols in the *varnamala* are learned,
the traditional method was replaced by books that introduced whole words from the first day.

The reading material is organized around a subset of Devanagari symbols. The *Introduction*
(Eklavya, 2003) explains that the symbols were selected based on their frequency of
occurrence: symbols that occur most frequently in children’s texts, such as stories, are taught
first. Explicit teaching of these symbols occurs in two stages:

Stage 1. क, व, स, म, प, न, ल, ल (These symbols represent the following sounds: k, b, s, m,
   p, n, l, and the diacritic for ā.)

Stage 2. र, घ, त, च, घ, थ, ध, ध, ध (These symbols represent the following sounds: r, g, t, tʃ,
   ḍ, and the diacritics for e and i.)

Note that these symbols have simple shapes. Digraphs, such as ख (kʰ) and श (sh) are not
included, and symbols that are frequently confused, such as घ and ध, are not taught together.
The selected symbols are also useful in forming words. The independent vowels, which are
the first symbols of the *varnamala* sequence, have been omitted because they usually occur in
word-initial cases, such as अब (ab meaning now). Instead, the more frequent and useful form
—the diacritic—is taught.

Figure 2 shows the first page of the textbook. Children begin by reading whole words. Familiar
words that can be represented in pictures are selected. Symbol recognition focuses on the
selected symbol set, but words with unfamiliar symbols and diacritics are also included to lend
authenticity to the texts.
Whole-word reading is made possible because this method takes advantage of the inherent schwa in every consonant. The method also uses a second property that is specific to Hindi—at the end of a word, the schwa is automatically deleted. Together these two properties permit learners to combine two symbols into words, such as nal (tap), cup (cup) and bus (bus) without learning additional symbols for the vowels. Students can learn to map letters to sounds (grapheme-phoneme correspondence) as well as manipulate symbols to create new words. For example, if they learn the symbols in the word bus, namely, बस, they can reverse the symbols to create a new word, which is सब (every). Note the close similarity to the method used by Mace (1961) to teach the Persian script.

Eklavya has not conducted any longitudinal or comparative studies to assess the effectiveness of this method. With a little imagination, this approach can be used with heritage learners to make script instruction more interesting and relevant right from the start.

6. Conclusion

The approaches described in this paper focus only on the first stage of learning an Indic script, namely, recognition of symbols. However, even at this stage, learners face problems discriminating and remembering the symbols. Although the traditional approach appears highly systematic to the linguist, it has not been designed for learners. Most learners accept that there is no way out of memorizing the symbols, but others drop out of language and literacy
programs if they fail to learn the symbols or if they see no relevance in the activity. The two methods outlined here approach teaching and learning from the perspective of the learner by identifying actual difficulties in symbol discrimination and learner concerns about relevance. In some approaches, a combination of these methods is used (for Telugu, see Sivarama Murthy, 2004).

Although these approaches have been in use for about two decades, no systematic comparative or long-term studies have been conducted. In fact, it is difficult to even obtain documentation about the different approaches used to teach Indic scripts because as one researcher put it, “We are too busy designing material to write reports on the results.” This poses a major problem in assessing the effectiveness of these approaches vis-à-vis the traditional approach that has been in use for centuries. It is important to compare the effects of these approaches with the traditional approach, with each other, and with other innovative programs that have not been documented here. In addition, different populations may benefit more from one approach than another—what does not work with illiterate adults may be effective with heritage language learners, and an approach that suits children learning their first script may not be as effective with foreign language learners. If we want to make our teaching more effective, we need to re-examine our traditional teaching methods and better align them with learner needs.

Acknowledgments

This paper grew out of discussions with Jennifer Bayer at the Central Institute of Indian Languages in Mysore and Anjali Noronha at Eklavya in Bhopal. I want to thank Brinda Ramesh, who arranged school visits in Ranikhet, and the school Principals—Dr. Rautela, Jyoti Shah, and Pauly Samuels—who gave me access to the classrooms and requested teachers to share their views with me.

References


There is no consensus on a name for the writing system that includes scripts such as Devanagari. These have been called Brahmi-derived scripts, semi-alphabetic scripts (Vaid & Gupta, 2002), semi-syllabic scripts (Vaid & Padakannaya, 2004), alphasyllabaries (Bright, 1999), and abugidas (Daniels, 1992, 1996). This paper uses the general term "Indic scripts".

A linguist’s approach to teaching the symbols can be seen in Veatch’s proposal, 'Literacy on a Sheet of Paper' at http://sprex.com/literacy/

When I was learning the Japanese syllabaries, I had no problem reciting the symbol sequence which is an abbreviated version of the varnamala, but I confused visually similar symbols, such as き (ki) and さ (sa). My Japanese teacher could not understand my confusion because to her the symbols came from different sections of the chart. I eventually learned Hiragana through the Kumon method which teaches the symbols through shape similarity.