Learning to read languages of South Asia

Sonali Nag

The Promise Foundation (India)

A later version of this paper appears as a book chapter in
Reading Acquisition Across Languages and Writing Systems
edited by
L. Verhoeven and C. Perfetti, Cambridge University Press. (published October 2017)

Introduction

South Asia is rich in linguistic diversity and a robust body of research is available to describe the distinct features of the multiple languages of the region (e.g., Dravidian languages, Krishnamurti, 2003; Indo-Aryan languages, Masica, 1991). Many South Asian languages share a common early history and have experienced similar waves of language contact over the centuries, leading to a phenomenon of convergence in lexicon, grammar and writing system. The overarching presence of such commonalities led to the region being called a linguistic area (Sprachbund, Emeneau, 1956). Acquisition research is available in several languages of South Asia but the body of evidence is small in each language. In this paper, cross-linguistic evidence will be used to explore literacy development in the Indic writing system. It is also critical to note that socioeconomic inequalities run deep in the South Asian region. Socio-political and socio-linguistic issues are therefore inextricably woven into any discussion on literacy development and schooling in the region (c.f. Sen, 2010). While some of the issues involved will be highlighted, my focus in this paper is on the psycholinguistic and cognitive-behavioural aspects of literacy learning.

1 sonalinag@t-p-f.org
1.1 South Asian languages and their writing system

The orthographies of a large number of languages of South Asia are descendants of the ancient Indic writing system called Brahmi. There are other writing systems in the region including the Arabic-derived scripts of Urdu and Pashto and the Latin-based script of Khasi, but these are not covered here.

Akshara is a basic symbol unit of the Indic writing system. While many akshara are constructed from phonemic markers (alphabet-like), some are representations of syllables with no discernible phonemic markers (syllabary-like). A typological description for the writing system is the *alphasyllabary* (Bright, 1996; Salomon, 2000) although it would be erroneous to infer from the term that the akshara system is merely a hybrid with part-solutions borrowed from an alphabet and a syllabary (a point also made by Bright, 1996, and Salomon, 2000). Other names for this writing system include syllabooalphabetic (Patel & Soper, 1987), syllabic alphabet (Coulmas, 1996), abugida (Daniels, 1996) and sub-syllabic (Vaid & Padakannaya, 2004). Figure 1 (column 1) gives a selection of akshara in four South Asian languages.

The akshara is minimally a vowel (/V/), lengthened by combining pre-vocalic consonants (/CV/, /CCV/, /CCCV/). A sub-set of symbols carries an inherent vowel (/Ca/) which has been said to offer articulatory ease to spoken consonants (Patel, 1996, 2004, 2007; Pandey, 2007a). Depending on the language, the inherent vowel is a mid-central [ə], open-central [a] or open-mid back rounded [ɔ] vowel. The automatically assumed inherent vowel in a consonant may be nullified by adding a vowel-suppression marker or by constructing conjoint consonants where a consonant gives up the inherent vowel by ligaturing to another consonant (see akshara marked as 1, 6 and 7 in Figure 1). The former is only sparingly used in some languages. Taken together, the phenomenon of inherent vowel and ‘consonantal conjoints’ are considered as defining features of the writing system (Salomon, 2000). Finally, a less transparent instance of vowel suppression is also available where a /Ca/ akshara
represents a vowel-less consonant $<C^0>$, with the phonological context needed for this inference (e.g. from Hindi: coda /r/ written with inherent vowel (/ra/) but read as /karna/, to do).

A principle of surface organisation moderates transcription of the spoken language such that markers for sound units fall into a series of block-like arrangements. Each vowel and consonant has a primary and a secondary form (e.g., in Hindi: /i/ = C, /i/; /k/ = M ç, Y), and positional rules govern their use. For vowels, the primary form is typically valid in word-initial position and for morphemic vowels, with the secondary form reserved for the post-consonantal position. Among consonants, the primary form is for syllable-initial consonants and the secondary form for consonants in a cluster. The visuo-spatial arrangement of some akshara may be non-linear, with phoneme markers placed in one of the four quarters of the symbol space (e.g., vowel markers in Hindi: ि, ी, ै, ौ, /i/, /ī/, /e/, /u/).

A principle of complete phonemic transcription governs written representation of spoken words (Figure 1, last column). For this, the rule is to represent the coda (syllable-final consonant(s)) as the next akshara. This is a re-syllabification phenomenon, carried intact from ancestral forms of the writing system into most contemporary varieties (Patel, 2007; Salomon, 2000; but see Miller, 2013).

Resyllabification for purposes of writing maintains the surface arrangement of symbol blocks. The coda has one of three possible representations: (i) a /Ca/ akshara with the inherent vowel assumed as suppressed (ii) an akshara with the vowel suppression marker, and (iii) an akshara that also carries information from the next syllable as either a coda-body or coda-syllable mapping (illustrated in Figure 1). A /CV.CV/ word may thus be written as $<CV.C^0.CV>$, $<CV.C.CV>$ or $<CV.CCV>$, and among these, it is the context-sensitive interpretation of the /Ca/ as $<C^0>$ that introduces opacity in an otherwise transparent system. The choice of the appropriate coda representation depends on rules of syllabification of the specific South Asian language, and the etymological, morphological and morphophonological characteristics of individual words. The forward-pinning of codas to the next phonological
syllable may form akshara representing a coda-syllable mapping or a coda-body mapping. The mapping of the writing system to spoken language is therefore at multiple levels (Figure 1). Consequently, the construct of orthographic depth (Frost, Katz, & Bentin, 1987) is to be analysed at two levels in the akshara writing system. Within the akshara, transparency is bound by positional rules for use of primary and secondary forms. Within a word, accommodation of re-syllabification may be transparent options (<CCV>, <C>), or the opaque option of a /Ca/ akshara with suppression of the inherent vowel.

**Figure 1:** Schematic representation of the akshara writing system: orthographic examples and levels of mapping to language.

In summary, it is useful to think of the akshara for what it represents when single and when embedded within a word. As a singleton akshara, the unit is an orthographic syllable with the exception of the phonemic consonant. In a word context, the akshara additionally may represent re-syllabification.
of the spoken string. We see below that insights about the links between the linguistic and orthographic layers are critical for literacy development in this writing system.

2. Acquisition of reading and spelling in South Asian languages

Children bring to the task of acquiring literacy a variety of cognitive and linguistic skills. It is interesting that writing system theorists have tended to invoke native-like proficiency to explain the learning of the akshara system. Thus learning, particularly those features that seemingly have no transparent linguistic rule, is seen to occur because of “the intuitive understanding of the native speaker” (Salomon, 2000, p 93), “lexical knowledge... is presumed” (Pandey, 2013, p 6), and because the reader is “simply expected to know” (Dasgupta & Sengupta, 2003, p 15). These explanations however are unsatisfactory, both from a theory-building point of view about the architecture and processing of a writing system, as well as for the practical concern of how children may best be supported to learn to read and write. The following sections try to untangle what these native-like proficiencies might be, while also showing that even among native learners of a writing system there are individual differences.

2.1 Becoming linguistically aware

2.1.1 Phonological development and phonological awareness

This section focuses on phonological development, examining intuitive syllabification and sub-syllabic processing skills. In intuitive syllabification, decisions about the position of syllable boundaries in a word can reflect phonological knowledge about the principles of sonority and legality. The akshara-based organisation of sounds in orthography may also mediate to align syllabification with the orthographic syllables rather than phonological syllables, with the nature of coda representation in particular exerting influence. Among Hindi adult readers this was indeed found to be the case; if \(<CV.C^0.CV>\) then segmentation was /CVC.CV/, but if \(<CV.CCV>\) then /CV.CCV/ (M. Ohala, 1999). Similar mapping of syllable segmentation to orthographic organisation is also reported in Malayalam (Mohanan,
Learning to read languages of South Asia

1989) while in Telugu the evidence comes from the selective treatment of the nasal coda called anuswara, which unlike the <CV.CCV> representations of all other codas, is an orthographically distinct marker without the forward pinning to form a /CCV/ akshara (Murty, Otake, & Cutler, 2007; Sailaja, 2007). Turning to intuitive syllabification in children, the trends are far from clear. In Bengali, intuitive syllabification of CVCCV nonwords by six-to-ten year olds showed a preference for a single consonant onset (/CVC.CV/) overriding both sonority and legality principles (Sircar & Nag, 2013a). Inferences about the influence of akshara-based representations could not be reliably made in this study because in Bengali both <CV.C⁰.CV> and <CV.CCV> representations are available for the coda and it was not clear which one might have been invoked by the child for the aurally presented nonword stimuli. This ambiguity is not there in Kannada where <CV.CCV> representation is the norm. A survey with five-to-eight year olds found two distinct strategies on CVCCV nonwords. For 54% of the sample, segmentations were essentially akshara-by-akshara (/CV.CCV/), suggesting a large influence of orthography on spoken language processing. But for the rest, a /CVC.CV/ syllabification was selective to certain items and structured around Kannada phonotactics, suggesting that these children ignored the orthographic layer. More research is clearly needed to understand influence of literacy in the akshara systems on syllabic processing. It is plausible that greater akshara knowledge increases akshara-based syllabification. But if better akshara knowledge is itself shaped by a greater understanding of the phonology of the language then it may also be that the mediation of orthographic organisation reduces with greater akshara knowledge.

Turning to sub-syllabic processing, the Kannada language is replete with word pairs minimally distinct at the phonemic level (‘āne’ – ‘āme’ (elephant-tortoise), ‘bēga’ – ‘bīga’ (fast – lock), ‘kūgi’ – ‘kūgu’ (verb forms for crying out), and morpho-phonological rules that introduce phonemic alterations at morpheme boundaries. Despite this demand in the spoken language for implicit phonemic
awareness, children show a slow pace of attainments on tasks requiring explicit phonemic processing. Surveys by Nag and colleagues (Nag, 2007; Nag & Snowling, 2012, 2013), showed phonemic processing to be associated with increasing Kannada orthographic experience, with more fluent readers being better on a phoneme deletion task. The relationship appeared to be one of reciprocal influence between characteristics of the writing system and the fine-tuning of the phonological domain.

Importantly, a good 22% of a middle school cohort of eight-to-twelve-year-olds continued to be at floor in phonemic processing tasks. These findings confirm earlier indications in the literature of a spurt in explicit phonemic processing approximately three years after literacy instruction, with performance of some children remaining exceptionally low (Kannada: Prakash, Rekha, Nigam, & Karanth, 1993; Gujarati: Patel & Soper, 1987). It is important to note that children in all the above studies were in instructional contexts that treated the writing system more or less like a syllabary. Findings are strikingly different in literacy environments that actively demonstrate the use of phonemic markers, and thus arguably reduce the syllabary-like notion of the writing system. In one such instruction setting, children in Grades 2-3 scored well above chance on phoneme deletion tasks (Bengali: Sircar & Nag, 2013b). Clearly, there is a need to examine the quality of insights that literacy instruction provide about the akshara and how these interact with children’s phonological development.

2.1.2 Morphological development and morphological awareness

The two simple sentences in Box 1 show some of the ways that inflectional morphemes convey event semantics (who-did-what-to-whom) and other grammatical information in Kannada. In the first sentence, even if we do not know whether Raaji is a girl or a boy, the gender marker on the verb (last word) gives this information, while the last word (verb) in the second sentence carries the tense marker for when uncle played the radio. Systematic acquisition studies on Kannada’s morpho-syntax in school age children are not available (but for studies with pre-literate children see Devaki, 1991; Lidz, Gleitman,
& Gleitman, 2003). The relationship between development of morphological, morpho-syntactic, morpho-phonological (sandhi) knowledge and literacy may be a reciprocal one. Seeing in print the morpho-phonological alteration in word boundaries, for example, may help to consolidate learning about sandhi changes in the index word and similar other words in the child’s lexicon. In addition, knowing many such instances may help to recognise these words more quickly in print. There may even be links between reading and spelling of multi-morphemic words, and an awareness of their morpho-syntactic functions. A preliminary indication that this may be so is the better performance found on a Kannada grammaticality judgement task among in-school children when compared to a never-enrolled group (Karanth, Kudva, & Vijayan, 1995). There is also evidence that four to seven year old children with more exposure to Kannada oral narratives and book reading at home are more accurate with morpho-syntactic details on a repetition task of just-heard sentences of differing syntactic complexity (Nag, Mirković, & Snowling, 2012).

2.2 Development of word identification

Two issues about akshara learning have implications for word identification. First is the obvious: words become easy to identify when individual akshara are accurately recognised. The factors influencing the learning of the full complement of akshara are several, some of which have only recently begun to receive research attention. In the perceptual domain, visuo-spatial factors, particularly non-linear arrangements, influence processing (Vaid & Gupta, 2002; Wali, Sproat, Padakannaya, & Bhuvaneshwari, 2009). In the experiential domain, frequency of occurrence in child-directed texts varies (e.g., Patel et al., 2013), and along with sequence and scope of instruction, influences pace of learning (Nag, 2007, 2011, 2013). Some of the key cognitive-linguistic factors that matter at the onset of literacy learning are vocabulary, visual memory, phonological processing skills, phonological memory and rapid automatized naming (RAN) (Kannada: Nag & Snowling, 2012; Sinhala: Wijayathilake & Parrila, 2013).
The last perhaps taps the efficiency of cross-modal learning between the visual and verbal domains.

During the primary school years, a reciprocal relationship between fine-grained phonological processing and greater analytical skills for phonemic markers appears to define the akshara learning system (see Nag, 2007; 2011; Nag & Snowling, 2012, 2013). By middle school, a marker for a poor reader is a child who struggles to recognize less frequent akshara or akshara with two or more phoneme markers (Nag, 2011; Nag & Snowling, 2011b).

The cognitive-linguistic process of accurate mapping of akshara to oral language also aids word identification. Recall that in this writing system, a guiding principle for transcription is to represent all sound segments, and that the principle of surface organisation into symbol blocks moderates this process. Akshara in the context of words therefore may map on to phonological syllables, the body or coda, and successive syllables either as coda-body or coda-full-syllable concatenations (see Figure 1). In Kannada, as is the case with other languages of South Asia that allow closed syllables, the resyllabification phenomenon of forward pinning of the coda to the next orthographic syllable (\(<CV.CCV>\)) implies that the /CCV/ akshara is exceptionally productive. It is therefore understood that even advanced readers will encounter unfamiliar /CCV/ akshara in the course of reading literary texts or because unusual /CCV/ akshara have been put together to transcribe foreign words (e.g., ‘phosphate’, ‘acquire’). It would also appear therefore that both insights into the principles of the writing system and recognising cues from the lexical context would aid word identification.

### 2.2.1 Word decoding development

A growing body of research shows moderate to strong associations between development of Kannada word decoding skills, akshara knowledge and phonological skills, and these trends appear in other languages of South Asia as well. Grades 1 to 4 are a period of actively expanding skills for decoding words (Nag, 2007, Prema, 1998). The challenge is particularly with more complex akshara in
words, as also reported in Bengali (Sircar & Nag, 2013b), Sinhala (Wijayathilake & Parrila, 2013), and Gujarati (Patel, 2004). Over these years, akshara knowledge expands and sub-syllabic processing improves both in Kannada (Nag, 2007; Prakash et al., 1993) and other languages (e.g., Patel, 2004; Tiwari, Nair & Krishnan, 2011; Sircar & Nag, 2013b). Evidence from naming accuracy and naming latency tasks (Nag, 2007; Vaid & Gupta, 2002) and parsing tasks (Patel, 2004) suggests that an analytic approach to word identification underpins this development. Possibly, awareness about the phonemic markers and multiple levels of mapping to phonology are factors that support children moving from being global to more analytic and strategic in their processing of words.

Further insights into word identification processes come from an analysis of component skills of reading accuracy and reading rate (fluency, where speed and accuracy are both expected). Nag & Snowling (2012), examining performance of middle schoolers, found that both syllable and phoneme level performance explained individual differences in Kannada reading accuracy, whereas for reading rate, only phoneme level scores were significant. It would appear that the orthographic properties of syllabic symbols and phonemic markers maintain sensitivity of both syllable and phoneme level processing. When the child cracks the akshara code and gains alphasyllabic competence, the biggest gains are with words that carry complex akshara. RAN was also a unique predictor for both components which is perhaps because RAN is related to well-automatized processes for access and retrieval from the extensive symbol set.

Another important factor in word identification is oral language beyond phonology. Strong associations between phrase repetition and word identification in phrases by 4-7-year-olds point to the possible links between spoken and written language (Ramachandra & Karanth, 2007). However, among older readers (15-45-year-olds), unusual akshara-phonology mapping and not semantic features of concreteness-abstractness, influenced naming latencies (Karanth, Mathew & Kurien, 2004) suggesting a disproportionately larger role of orthographic processing in Kannada word identification. In other South
Asian languages, individual differences have been found in primary schoolers’ accuracy on word level parsing tasks (e.g., Gujarati: Patel, 2004; Sinhala: Wijayathilake & Parrila, 2013). Here again it is difficult to draw any firm conclusions about the role of morphological processing in word identification because all words are treated as similar and no morphological analysis is reported. Nonetheless, using one’s lexical repertoire is a valuable strategy particularly since akshara learning takes time to reach a high level of efficiency. Moreover, wider lexical knowledge, including implicit knowledge of phonotactics, etymological conventions, and morpho-phonological rules is a valuable resource when akshara-phonology (and phonology-akshara) mappings are ambiguous. One preliminary line of evidence for this proposal comes from a Bengali nonword reading task. Sircar and Nag (2013b) exploited the nature of coda representation in Bengali - the medial akshara in <CVCaCV> may map either to /CVCaCV/ or /CVC⁰CV/ - to examine decoding strategies. Phonological analogies in the lexicon appeared to aid decoding at points of ambiguity. Thus the nonword <pīraṣā> was decoded by typically developing readers in Grades 3-4 as a /CVCaCV/ in keeping with words like ‘Dhāranā’ (assumption) and ‘preraṇā’ (inspiration), while <DaTakā> was decoded as /CVC⁰CV/ in keeping with common words like ‘maTka’ (pot) and ‘paTka’ (firecrackers). Importantly, while typically developing readers made a distinction between the two nonwords, less skilled and younger readers were unquestioning of orthography, reading both nonwords as /CVCaCV/. Quality lexical representations, it appears, makes for a more discerning identification of written words.

2.2.2 Word spelling development

A bidirectional mapping between phonology-akshara and akshara-phonology is involved in spelling. Accurate phonology-akshara mapping put down in the right sequence of orthographic syllables best describes the process of spelling in South Asian languages. Reports of spelling of primary and middle school children show individual differences in accuracy at the level of mapping as well as selection of
orthographic syllables (Hindi: Vaid & Gupta, (2002); Gujarati: Patel, (2004), Telugu: Vasanta (2004)). A study by Nag, Treiman, & Snowling (2010) of the Kannada spellings of middle schoolers, captures many of the key issues in learning the component skills of this process. Spelling shorter words, with only akshara that have the inherent vowel (/Ca/) is easier than longer words and words that need attention to multiple phonemic markers. Consonant segments within akshara are usually easier to spell, though phonologically close neighbours (e.g., the fricatives /s/, /ṣ/, /ś/) are prone to error, and, as in reading, consonants in the /CCV/ akshara need particular attention. Common errors with vowels are a) omitting to write their phonemic markers (appearing then to revert to the inherent vowel), and b) a phonological mix-up within the short-long vowel pair (e.g., /i/ for /ī/).

Moving next to factors in akshara-phonology mapping that influence spelling, Nag (2013) found that when words were matched for age of acquisition and frequency in child-directed print, children in Grade 3 did better on those akshara that represented a phonological syllable (/dʰya/ in /dʰya.na/) than those that represented successive syllables (/bda/ in /sʰab.da/). There were large individual differences in children’s awareness of the principles of akshara-phonology mapping, making it clear that the cognitive-linguistic underpinnings of akshara-based spelling is more than mere recall of individual symbols.

Two socio-linguistic processes also influence spelling in Kannada, and other languages of South Asia (e.g., Sugatapala De Silva, 1986; Patnaik & Pandit, 1986). First is diglossia (e.g., the dropping of the glottal fricative /h/ in certain dialects of Kannada so /halva/ spelled as <alva>; Nag et al., 2010). Second is the process of nativising loan words, and confusion about phonological and morpho-phonological alterations that apply. Examples of spelling versions of English words in Kannada are television (‘televijan’, ‘televishan’), ticket + plural marker (‘ticket.tannu’, ‘ticketu.galannu’) and tender + plural marker (‘tendaru.galannu’, ‘tendar.galannu’ (Ranganatha, 1982).
2.2.3 Reading and spelling difficulties

A current view of reading difficulties is that poor reading can be caused by an accumulation of cognitive risk factors with some cognitive factors also acting as protective influences that moderate the impact of a risk factor (Snowling, 2008). Findings from the few available studies on children with atypical literacy development confirm such a multi-factorial explanation for the reading and spelling difficulties in the languages of South Asia. Cognitive profiles of poor readers show that single deficit profiles are rare and lower attainments are recorded in symbol learning and sub-syllabic processing (Hindi: Gupta, 2004; Kannada: Nag & Snowling, 2011b; Rama, Miles, & Lalithamma, 1993), as well as knowledge of vocabulary and awareness of the morpho-syntax of the language (Nag & Snowling, 2011b). Reduced speed during rapid automatized naming is also reported for children in the lower end of the attainments distribution (Nag & Snowling, 2012; for Sinhala see Wijayathilake & Parrila, 2013). The available data suggests two hypotheses about dyslexia in the akshara-based languages: one related to a deficit in phonological processing and the other to a mapping problem which could in principle stem from the phonological deficit (Snowling & Nag, 2011). Figure 2 gives a summary of the type of deficits that might characterise each. Within this profile of deficits, it is possible that high quality semantic and morpho-syntactic knowledge may act as a protective factor. Systematic case studies as well as group comparison studies are needed to test these hypotheses.

Figure 2: Types of deficits related to two hypotheses about Dyslexia in South Asian languages
2.3 Reading comprehension

Reading comprehension surveys in the languages of South Asia report three reading profiles (e.g., Gujarati: Patel, 2004; Bengali: Nag & Sircar, 2008; Kannada: Nag, 2008). One group of children reads fluently and is typically accurate on factual questions about the text, a second group reads *akshara-by-akshara*, but seems to gather meaning and records average reading comprehension scores, and a third group which fails in both decoding and reading comprehension. It would appear that the quality of orthographic representation of words in these sub-groups is different. Also perhaps, children vary in the extent to which they draw upon lexical representations to disambiguate written words and therefore scaffold comprehension. Extending such a proposal to a study of middle school reading, Nag & Snowling (2012) predicted that higher attainments in Kannada reading comprehension would be associated with higher lexical awareness. The latter was assessed through a vocabulary definition task, and a test of accuracy when recalling the content of richly inflected sentences. By contrast, it was predicted that those with lower levels of reading comprehension would show diminished vocabulary knowledge and mapping of morpho-syntactic information to meaning. Analysis of data from 95 children in 12 schools in Grades 4–6 confirmed these predictions: significant positive correlations were found between reading comprehension and vocabulary (.340), and inflection knowledge (.354). Strong correlations were also found between reading comprehension and accuracy (.644), and phonological processing (.594). Overall, the children who were better at reading comprehension were also better in reading accuracy, phonological processing, knowledge of vocabulary and inflectional morphology.

2.3.1 Predictors of reading comprehension

Turning to concurrent predictors in the same group of Kannada speaking children, reading accuracy and phonological processing measures together accounted for 44% of variance in reading comprehension, after controlling for age. In a third step of the analysis, children’s vocabulary scores or
scores on the inflection knowledge task were entered separately. Each measure predicted unique variance in reading comprehension: vocabulary predicted 4.8% and inflection knowledge 2.2% (the non-verbal general abilities measure was excluded from the Model because it did not show associations with reading). This Model gives the cognitive-linguistic underpinnings of reading comprehension in Kannada, and is perhaps as relevant for the other akshara-based languages of South Asia. First, since reading comprehension is essentially dependent on accurate word identification, and some aspects of akshara knowledge are still being mastered in the middle school years (Nag, 2007, 2013), it is not surprising that a substantial proportion of the variance in reading comprehension was explained by skills for decoding and phonological processing. Second, vocabulary and grammar knowledge support the generation of meaning from the text, and therefore, again as predicted, these aspects of broader oral language explained variance in reading comprehension.

Inference making is another factor contributing to individual differences in reading comprehension. Small scale surveys show many primary schoolers making connections between idea units to draw simple inferences, but there is also a lot of variation in attainments on this skill (Bengali: Nag & Sircar, 2008; Kannada: Nag, 2008). It is important to note that the instruction methods available to the children in these surveys did not actively encourage inference making, a contextual factor that is quite widespread in the South Asian region (Gujarati: Patel, 2004; Dyer, 2008; Tamil: Geeta, 2012; Hindi: Jayaram, 2008a, 2008b). Research that disentangles instruction effects from spontaneous individual differences in inference making is needed to throw light on reading comprehension processes in such contexts. Also of interest is how phenomena such as causality, transitivity and other aspects of events are encoded in South Asian languages, and what this might mean for reading comprehension (e.g., ambiguities in causality in Kannada: Lidz et al., 2003; and ambiguities in word order in Hindi: Mishra, Pandey, & Srinivasan, 2010).
2.3.2 Word level effects in comprehending text

Words take root in the phonological, morpho-syntactic, morpho-phonology and orthographic domains and the more detailed their representation in each domain the higher their ‘lexical quality’ (Perfetti, 2007). Several strands of evidence in the preceding sections from Kannada and other languages of South Asia point to the influence that these multiple dimensions of lexical quality play in akshara-based reading. In addition, there is the domain of morpho-orthography, details of which are not yet available from child studies but preliminary evidence from skilled adult readers suggests word level effects (Hindi: Rao, Soni & Chaterjee-Singh, 2012).

2.4. Conclusion

There are reciprocal links between linguistic awareness and literacy development. Insights into akshara formation and improved automaticity in accessing and retrieving mapping details when reading and spelling are intricately and reciprocally associated with phonological and morphological awareness. Further, literacy instruction has a far reaching effect from development of explicit phonemic awareness to inference making during text reading.

3. Discussion

3.1 Challenges in learning to read South Asian languages

The akshara orthographies are almost as productive as the Chinese sets, with one crucial difference – a combinatorial principal dramatically reduces the number of akshara to be individually learnt. Thus, if the consonant is fixed as the base for constructing the symbol block, then the object of learning is reduced to a consonant-vowel matrix. But even then, the challenge of size remains because, as shown in the preceding sections, any number of akshara with consonant clusters must be constructed and most of these are not explicitly taught. Learning hurdles also come from the visual complexity of the symbols,
their phonological confusability and mis-sequences in visuo-spatial arrangements. A final challenge is about knowing which phonological units in a word an akshara encodes, and vice-versa. Linguistic knowledge about syllabification, etymology and morphology provides the insights for gaining mastery over these orthographic complexities.

3.2 Implications for instruction

This section discusses issues related to recitation, copywriting, and reading prescribed books since these activities dominate many early literacy classrooms in South Asia, especially in public-funded schools (Dyer, 2008; Geeta, 2012; Jayaram, 2008a, 2008b; Patel, 2004; Nag, 2007; Vagh, 2010).

Recitation and copywriting. An efficiently laid out synthetic phonics scheme provides the material for sound drills in all akshara-based languages of South Asia: Among consonants, the unvoiced and voiced sounds, and unaspirated-aspirated pairs are laid out in strings (e.g., /ka, kha, ga, gha/, /pa, pha, ba, bha/). The confusing nasals (variations of /n/) are given prominence individually at the end of each phonics string. The fricatives (variations on /s/) come in a sequence, with the trills (/r/), liquids (variations on /l/) and aspirants (like /y/ and /v/) interspersed for maximum perceptual distinctiveness. Short and long vowel pairs (/a – ā, i – ī/) make another set. Among the close and close-mid vowels, perceptual distinctiveness is enhanced by interleaving front vowels with back vowels (thus /i-u-e-o/). A final akshara set, reserved for Grades 2-3 usually, has a small collection of high frequency consonant clusters. The potential of such phonics materials for developing phonological awareness seems obvious, though it does not contribute to explicit phonological processing at the sub-syllabic level.

Akshara charts are prominently displayed in classrooms and are an easy-to-refer visual toolkit for the beginning reader. Copywriting of these akshara is common. The parallel practice of copywriting and recitation arguably draw attention to fine-grained visuo-spatial details and their linkages to sound. But since no explicit attention is brought to the phonemic markers in these de-contextualised lists of
symbols, the accompanying sound drills can promote a syllabary-like notion of the writing system. Incidental learning then becomes important for children to extract the underlying rules of akshara formation.

The available research does not provide a clear account of the learning mechanisms targeted by these indigenous methods of literacy instruction though there are several pointers to make them more efficient: First, the traditional phonics charts described above do not have to be an invariant sequence for teaching about the akshara. A better starting point is instead with high frequency akshara, particularly those akshara that help construct words common in early vocabulary. Second, explicit instruction of parts of an akshara help children abstract combinatorial rules, including copywriting of the same phonemic marker (e.g., the vowel markers) with the whole complement of consonants. Third, the separation of symbol sets into /V/, /Ca/, /CV/ and /CCV/ is artificial and does not reflect the phonological patterns of the spoken language. Children learn about all of these potentially easily-recognisable akshara from encounters within known words, and hence there is no need to sanitize early grade readers of all complex akshara (like /mma/ in /amma/, mother). Finally, akshara charts do not address the mapping issue. What appears to strengthen insights is good quality lexical representation of words, and a large vocabulary. Hence, a robust oral language programme must accompany akshara practice, even when the language is the home language of the child.

*Reading books.* If akshara learning is frequency-sensitive then the more words in which they come embedded, the greater the chance of quickly gaining mastery. Thus, while reading practice with books is growing in most South Asian classrooms, the value of the practice needs to be qualified – repeated reading of the same book (e.g., one prescribed textbook) severely limits opportunities for implicit akshara learning. Similarly, artificially graded books with restricted vocabularies cannot offer the variety needed for akshara learning. Moving to higher order engagement with texts, the evidence suggests that reduced vocabulary and a grasp of the grammar of the language slow down attainments.
Taken together, an important implication from akshara literacy research is to introduce children of all ages to variety and complexity in narratives, at the spoken and written level.

4. End note

Literacy research in the languages of South Asia is mainly descriptive giving information of associations between variables. The available evidence gives more information on what-goes-with-what during literacy learning in this writing system. A very small set of studies are with sequential and process data that can give an account of ‘what comes from what’ or ‘what leads to what’. From the evidence available, it appears that literacy development moves from a global view of the akshara to an analytic understanding. This transition is possibly driven by a growing awareness about component phonemic markers and their mapping to multiple levels of phonology. Morphological awareness and akshara knowledge related to resyllabification are further factors that support word-level processing. Turning to atypical development, hypotheses under consideration for literacy difficulties are a phonological processing deficit, and a deficit in the mapping principle. Difficulties with vocabulary and grammar learning are additionally common among poor readers, with the intersecting domains of morpho-phonology and morpho-orthography yet to be explored. Taken together, the picture is clear that gaining mastery in the Indic writing system requires insights into the intricate mapping between the spoken language and its written form, and knowledge about the complexities of its vocabulary and grammar.

For complete reference list contact the author. sonalinag@t-p-f.org