Learning to spell in an alphasyllabary: The case of Kannada

Sonali Nag a, Rebecca Treiman b & Margaret J. Snowling c

a The Promise Foundation, India and Centre for Reading and Language, University of York, UK
b Department of Psychology, Washington University in St. Louis, USA
c Centre for Reading and Language, University of York, UK

Learning to spell in an alphasyllabary: The case of Kannada

Sonali Nag
The Promise Foundation, India and Centre for Reading and Language, University of York, UK

Rebecca Treiman
Department of Psychology, Washington University in St. Louis, USA

Margaret J. Snowling
Centre for Reading and Language, University of York, UK

Abstract
We investigated children’s spelling in the alphasyllabic writing system of Kannada, a South Indian language. Kannada represents language at roughly the level of the syllable, but its symbols or ‘akshara’ can be segmented visually to reveal the consonant and vowel segments within. Having been taught akshara as whole, fourth- and fifth-grade children were poor at constructing complex akshara from their consonant and vowel segments. Analyses of their errors showed confusions between the symbols for phonologically similar consonants, visually based errors, particularly on small diacritic marks, and influences of dialect. Kannada has been considered a transparent writing system, but the challenges that arise in mastering this visuo-spatially complex and extensive orthography suggest that previous studies have missed certain factors that contribute to the ease or difficulty with which a writing system is learned.

1 Introduction
All writing systems represent language, but they differ in how they do so. Whereas alphabetic scripts represent speech at the level of the phoneme, syllabaries represent it at the level of syllables. ‘Alphasyllabaries’ (sometimes also called abugidas) have characteristics of both alphabets and syllabaries (Bright, 1996; Daniels, 1996; Sproat, 2006). These writing systems use graphic complexes that correspond roughly to syllables, but these symbols can be decomposed to show the phonemes within them. For example, the Kannada symbol के, or ‘ke’ in the Romanized form that we use at some points in this article, stands for /ke/. It is composed of a subpart that represents the phoneme /k/, ण, and a subpart that represents the phoneme /e/, ए; these subparts recur in other syllables that include /k/ and /e/. [The phoneme symbols we use are from the alphabet of the International Phonetic Association (IPA), 1999.] Although much research has been done on the acquisition of alphabetic writing systems, especially English, less is known about the acquisition of other systems. This situation has justifiably raised doubts...
about the relevance and explanatory power of theorizing that has emerged out of Anglo- and alphabet-centric research (e.g. Share, 2008). It is against this backdrop that we investigated one aspect of literacy development, spelling, in Kannada.

Alphasyllabaries might be thought to possess the advantages of both alphabets and syllabaries. One can spell a word one chunk at a time, eliminating the need for analysis into phonemes, or one can alternatively build up graphic complexes from their phonemic components. Indeed, Karanth (2006) reported that learners of Kannada have few spelling difficulties apart from those involving dialect features. However in another alphasyllabary, Tamil, Aaron and Joshi (2006) reported spelling errors that are independent of dialect features. Surveying the published research, we found no previous empirical investigations of spelling development in typically developing learners of the Kannada alphasyllabary. Before introducing our work on the topic, we describe some characteristics of the Kannada language and writing system.

Kannada is a South Dravidian language spoken in the state of Karnataka in South India. The majority of its words are bi- and trisyllabic, with four-, five- and six-syllable words also in the vocabulary. Monosyllabic words are rare in Standard Kannada but exist in some dialects. Standard Kannada has 43 phonemes (Upadhyaya, 1972). Velar plosives, palatal affricates, retroflex plosives, dental plosives, and bilabial plosives are distinguished in both aspiration and voicing. The consonant inventory also includes glides, liquids, fricatives, and nasals. Some consonants in each class differ only in place of articulation, including /l/ and /r/ for liquids, /n/ and /t/ for nasals, and /s/ and /z/ for fricatives. Standard Kannada has five vowels (/i/, /e/, /a/, /u/, /o/), each of which has a short and long form and two diphthongs. Syllables with single initial consonants are the most common type. Sequences of two consonants exist at the beginnings and in the middle of words, but those with more than two consonants are rare. Word-final syllables are typically open, and loan words are adapted to fit this rule.

The Kannada writing system, like a number of other writing systems used in India and nearby countries, has its roots in Brahmi, an ancient script of the Indian subcontinent. Kannada has over 400 of the written symbols that are called akshara, and other Indian writing systems contain between 200 and 500 symbols. An akshara is a graphic complex that represents, roughly speaking, a syllable, and all phonemes in the unit are represented in the symbol. In most akshara, as with ಫ for /ke/, the portions representing the different segments are physically joined to one another. In some akshara, as with ಸ for /kse/, one portion is separate from the rest. An akshara represents all contiguous consonants in a word, plus the following vowel. In some cases, as with ಆಗಸ 'washerman', the orthographic syllables match the phonological syllables. In other cases, as with ಹಲ್ವ 'a sweet dish', there is a mismatch between the orthographic syllables and the phonological syllables. Some postvocalic nasals and /r/ are an exception to the pattern just described in which all contiguous consonants are grouped with the following vowel. According to Karanth (2006), these exceptional patterns are one of the few aspects of written Kannada (apart from dialect features) to cause difficulty for spellers.

Each consonant and vowel may be represented in two main ways, with a primary symbol and with a secondary form. The primary symbol is written on the line of print, and the secondary form may be on the line, below it, or above it. The secondary form is usually smaller than the primary form. The secondary form for some consonants is a miniature version of the primary symbol, as in the case of ಙ (the secondary form for /k/) and ಙ (the primary symbol). The secondary forms of other consonants differ in shape from the primary form, as with ಡ (the secondary form for /m/) and ಌ (the primary form). For vowels, the secondary form always has a different shape from the primary form. For example, /a:/ has the primary form ಏ and the secondary form ಌ. The secondary forms of some long vowels are elaborations of the secondary forms of the corresponding short vowels. For example, the secondary form of /i:/ * is an elaboration of the secondary form of the corresponding short vowel /i:/ *. The secondary forms of vowels are often referred to as ‘diacritics’. In European scripts, diacritic marks are generally written above or below the basic letters and are often slightly separated from them. In contrast, the vowel diacritics of Kannada

S. Nag et al.
Kannada spelling

and other Indian scripts may be on the same line as the rest of the print, often physically joined to the base symbol.

Although each consonant has more than one possible spelling—a primary form or a secondary form—the choice between the two forms is governed by rules. An akshara is generally built with the consonant as the base, using the primary symbol for the consonant. The secondary forms for the other segments in the akshara are added to that base. When a consonant in a CV (C = consonant, V = vowel) is followed by a vowel other than the short ‘a’, the secondary vowel symbol is added to the primary consonant symbol. When a consonant is followed by the short vowel ‘a’, which in many cases is pronounced as /ʌ/ and which is called the inherent vowel, no vowel symbol is added to the primary consonant symbol. For example, the primary form of /t/ with the inherent vowel, त, is used to spell /ta/. The secondary form of the vowel /e/, इ, is added to the base consonant to spell /te/, ते. Secondary forms are used to spell consonants in the second and later positions of consonant sequences. The secondary form corresponding to the second consonant is typically located below the first consonant, and the secondary form corresponding to the vowel is added to the consonantal complex. For example, /t/ is written with the secondary form of /e/, इ, joined to the base consonant त, and the secondary form of /ʌ/, अ, added below. Consonants are more often spelled with the primary form than the secondary form because sequences of consonants are less common than CV sequences in the Kannada language. For vowels, on the other hand, the secondary form is more common than the primary form. The primary form of a vowel is used only when the vowel is at the beginning of a word or when it represents a separate morpheme.

There are some gaps between the spoken and the written form of Kannada. For example, speakers tend to delete or reduce many short vowels that are written in words. In the dialect of the area where the present study was conducted, the glottal fricative /h/ is sometimes dropped and final /e/ changes to /a/. Aspirated stops are not present among nonliterate people in this area, including nonliterate children. However, literate adults distinguish between aspirated and unaspirated forms (Manjulakshi, 1996).

In the schools where we carried out our research, formal Kannada literacy instruction begins in Grade 1, between the ages of 5 and 6 years. Charts showing sets of akshara are prominently displayed in classrooms and are used to teach the symbols. The primary vowels and consonants with the inherent vowel are introduced in Grade 1, and consonants with other vowels are taught in Grade 2. This is followed by instruction about a selection of frequently used CCV akshara. In this instruction scheme, the consonant with inherent vowel is the standard reference for a consonant symbol. Children are taught the akshara as wholes, and they are not explicitly told what the parts within each akshara represent. This approach is widespread in the region, and it is in some ways similar to approaches that are used in the teaching of alphabetic writing systems that focus on the memorization of whole syllables.

The foregoing description of the Kannada writing system and how it is taught suggests that learners of this system may encounter certain difficulties in spelling individual akshara. We will now describe past research that led us to expect spelling errors beyond those caused by dialect features and exceptional consonants, the two areas of difficulty that Karanth (2006) pointed out.

Research on the learning of alphabetic writing systems shows that children tend to spell better when they have been explicitly taught the correspondences between letters and phonemes than when teaching has focused on larger units (Ehri et al., 2001; Rayner et al., 2001). In alphasyllabaries, it is possible to teach spelling either at the akshara level or at the level of individual phonemes. Given the teaching methods that are used in the schools in the present study, it may be that children rely largely on rote memorization of whole akshara. If so, the large size of the akshara set may cause difficulties. Children who learn and retrieve akshara as whole symbols would be expected to make a large number of errors that involve substitutions with akshara that have visual similarities at a global level.

If learners of Kannada attempt to spell using a more analytic approach, working at the level of the phonemes in the akshara, a different pattern of errors would be expected. Children might make predominantly phonological errors, misspelling an akshara
or part of an akshara using a symbol that stands for a phonologically similar segment. They might have difficulty constructing complex akshara even when their forms are predictably based on knowledge of consonant and vowel segments and rules for combining them. Thus, even though the choice between the primary and secondary forms of the phonemes is governed by rules, children might have difficulty learning and using those rules. Consider the usually untaught English rule that should allow people to accurately choose between ‘l’ and ‘ll’ as a spelling of final /l/ in monosyllabic words. The rule states that ‘ll’ is used if the vowel is spelled with one letter, as in ‘wall’, and ‘l’ if the vowel is spelled with two letters, as in ‘wheel’. Despite the reliability of the rule, it takes some time for children to apply it correctly (Hayes et al., 2006).

The rules governing primary and secondary spellings of Kannada phonemes may be difficult to learn, in part, because the two spellings of a phoneme often differ in shape as well as size. This is somewhat similar to the situation with upper- and lowercase forms of the Latin alphabet, where pairs such as ‘R’ and ‘r’ differ in shape and pairs such as ‘C’ and ‘c’ differ only in size. Children in the USA have more difficulty with lower-case letters like ‘c’ than letters like ‘r’. Kannada learners are taught the primary forms of consonants and vowels before the secondary ones. For similar reasons, we might expect them to have difficulty with secondary forms that are different in shape from the better known primary forms.

Further difficulties may be caused by the fact that the secondary forms in Kannada are often small and not on the main line of print. Some alphabetic scripts also include forms of this kind. For example, the diacritic ´ may appear on ‘n’ in Slovak. Under some conditions, learners of Slovak have difficulty with such marks (Caravolas and Mikulajová, 2008). We anticipated similar problems with secondary forms in Kannada.

Given the considerations just described, we expected that even learners of Kannada who have had several years of formal literacy instruction would have some difficulties in spelling. We thus examined the spelling performance of 4th and 5th graders on a selection of words that included many of the features just described, asking whether the children’s errors reflect rote visual memorization of whole akshara or whether they arise in a process of constructing akshara from smaller phonological units.

2 Method

2.1 Participants

We analyzed spellings produced by 60 children (36 boys, 24 girls) in Grades 4 and 5, the last two years of primary school, in the Chamarajanagara district in South Karnataka, India. Consent was received from the education authority in accordance with the norms followed by the government school system in the country. Eighty percent of the children in the sample reported Kannada as their home language, and 60% were either bi- or multilingual. Since multilingualism is typical of India, we decided to study all children together rather than restrict ourselves to monolinguals.

We wished to examine typically developing spellers for the present study, and so we excluded children who were below the 15th percentile for their grade in overall accuracy in the spelling task. Some of the analyses further differentiated the sample along a continuum of spelling skill, comparing one group of children who had 15 or fewer correct responses on the 30-item spelling task to a second group with 16 or more correct responses. The former group, which we call the less-skilled spellers, included 24 4th graders and five 5th graders. The latter group, which we call the skilled spellers, included 10 4th graders and 21 5th graders. We broke down the results in this way, rather than by grade level, because there was a fair amount of variability within children in the same grade and some overlap across grades. Table 1 provides additional information about the skilled and less-skilled groups. The two groups did not differ significantly in age, $t(58) = 1.26$. 

---

2.2 Tests and materials

The children were tested four months after entry into the current grade. The spelling test was given to each child individually in a session that also included other tasks. The examiner spoke a Kannada dialect that was similar to that spoken in the area. The spelling list comprised 30 words of different types, which are explained below. The words, which are shown in the Appendix A together with their IPA forms and their translations, were taken from books written in Kannada for children in Grades 3–6. All of the words were expected to be familiar to all the children in their spoken form. There were 9 two-syllable words, 20 three-syllable words and 1 four-syllable word. There were no monosyllables because, as mentioned earlier, these are rare in Kannada and those that do exist were considered too easy for a spelling test for children of this age.

Given the lack of availability of word frequency or akshara frequency lists for Kannada, we were not able to select the words on that basis.

3 Results and Discussion

3.1 Word-level analyses

We examined spelling correctness on whole words as a function of the type of akshara they contained. Words of the first type, such as ಬಣ (bana ‘forest’) and ಪಡಾಕ (padaka ‘medal’) contained only CV akshara with the inherent vowel. Knowledge about secondary forms and how they are added to the base is not required for correct spelling of these words. Words of the second type contain at least one CV akshara with a vowel other than the inherent vowel but no consonant clusters, such as ಹಸಿರು (hasiru ‘green’) and ಬಾಗಿಲು (baagilu ‘door’). Words of the third type contain at least one CCV akshara. These could involve sequences of two identical consonants, as in ಮನಸು (manassu ‘mind’), or sequences of different consonants, as in ಹಲ್ವ (halva ‘a sweetmeat’). We will refer to the former sequences as ‘geminates’ and the latter sequences as ‘mixed clusters’. Akshara with a vowel other than the inherent vowel but no consonant clusters include one secondary form (secondary form /a:/ in /bala/) and akshara with a geminate consonant or consonant cluster have two (secondary forms /s/ and /u/ in /ssu/).

Table 2 shows the spelling accuracy of the two groups of children on the three types of words. A 3 (word type: only CV with inherent vowel, at least one CV without inherent vowel but no CCV, at least one CCV) × 2 (spelling skill: skilled, less skilled) within-subjects analysis of variance (ANOVA) revealed significant main effects of word type, \( F(1,58) = 311.56, \ p < 0.001, \ \eta^2 = 0.83 \) and skill, \( F(1,58) = 84.49, \ p < 0.001, \ \eta^2 = 0.59 \), as well as a significant interaction, \( F(1,58) = 6.31, \ p < 0.05, \ \eta^2 = 0.02 \). Children performed best on the words that contained only akshara with the inherent vowel, intermediate on the words that contained at least one CV akshara with another vowel, and poorest on the words that contained at least one cluster. The skilled spellers performed significantly better than the less-skilled spellers on all three types of words, but the difference between the groups was larger for the second and third categories of words—those that required knowledge of secondary forms—than for the first category of words, which did not require such knowledge. The children’s poor performance on words in the second and third categories suggests
that they did not know enough about vowel diacritics and the secondary forms of consonants to construct a number of the complex akshara.

A second analysis compared performance on bisyllabic words with performance on longer words. (A single analysis looking at both akshara type and word length could not be performed because the variables were not balanced in the word set.) A 2 (word length: bisyllabic, longer) × 2 (spelling skill: skilled, less skilled) within-subjects ANOVA found significant main effects of word length, $F(1,58)=105.09, p<0.001, \eta^2=0.64$ and skill, $F(1,58)=146.81, p<0.001, \eta^2=0.72$, but no interaction. For the less-skilled spellers, the mean proportion of correctly spelled words was 0.51 (SD = 0.09) for bisyllabic words and 0.35 (SD = 0.09) for longer words. For the skilled spellers, the figures were 0.76 (SD = 0.09) and 0.62 (SD = 0.13), respectively.

### 3.2 Akshara-level analysis

We next examined performance at the level of individual akshara. We examined three types of akshara: CVs with no consonant cluster and the inherent vowel, which include no secondary forms; CVs with vowels other than the inherent vowel, which have one secondary form; and CCVs, which have two secondary forms. Primary vowels and the exceptionally spelled consonant sequences were not included in this analysis because the list contained few such akshara. We distinguished between akshara in the first, second, and third position of a word, not including in this analysis the one akshara that appeared in the fourth position of a four-syllable word. Counting all akshara in the thirty-word spelling list except for the ones just mentioned, seventy-four items were included in the analysis. The first three rows of data in Table 3 show the proportion of akshara of each type spelled correctly by the two groups of children. Performance of the two groups on akshara in the first, second, and third positions in a word are shown in the last three rows of the table.

We conducted a regression analysis by items, using type of akshara, position of akshara in word, and length of word (bisyllabic versus trisyllabic) as predictors of spelling correctness. We did this separately for the skilled and less-skilled spellers. In the skilled group, akshara type was the only factor that contributed significantly to spelling performance ($\beta=0.62, p<0.001$). For the less-skilled group, spelling performance was predicted by both akshara type ($\beta=0.77, p<0.001$) and akshara position ($\beta=0.18, p<0.05$). Both groups performed more poorly on akshara that contained consonant clusters, or those that included two secondary forms, than on akshara that contained either one secondary form or no secondary forms. For the less-skilled spellers only, akshara in the third position of words were significantly more difficult to spell than those in the first or second positions. Word length did not make a reliable contribution once other variables were held constant. For example, children did not spell an akshara better when it was the first syllable of a bisyllabic word than when it was the first syllable of a longer word.

The results of the akshara-level analysis confirm the results of the word-level analyses in showing that akshara that include secondary forms, especially those with two such forms, are difficult for 4th and 5th graders to spell. For less-skilled spellers, there is an additional position effect such that akshara later in a word are harder than those earlier in the word. The less skilled spellers may be expending more effort on the early parts of the word, leaving little attention available for the final portion.

### 3.3 Intra-akshara analyses

So far, we have considered correctness at the level of whole words or whole akshara. We now examine
children’s performance on the components of individual akshara. We began by comparing performance on consonants and vowels on the 60 CV akshara. Any substitution on the consonant segment was coded as a consonant error. For akshara with the inherent vowel, the addition of a vowel diacritic was coded as a vowel error, and for akshara with other vowels a vowel error was coded if an incorrect vowel diacritic was used or if the required vowel diacritic was omitted. A 2 × 2 within-subjects ANOVA using the factors of segment type (consonant, vowel) and skill (skilled, less skilled) found significant main effects of type, \( F(1,58) = 28.76, p < 0.001, \eta^2 = 0.33 \) and skill, \( F(1,58) = 20.36, p < 0.001, \eta^2 = 0.25 \), but no interaction. Both groups of students spelled consonants better than vowels. The mean proportions of correct spellings on consonants and vowels were 0.93 as compared with 0.87, respectively, for the less-skilled group and 0.97 as compared with 0.91 correct for the skilled group.

To examine the nature of children’s consonant errors, we categorized substitution errors on consonants along two features: visual and phonetic. A substitution was considered visual when the substituted symbol had either one more or one less visual feature than the correct symbol, with a line or a curve in the akshara being considered as a visual feature. We considered a substitution phonological if the substituted symbol belonged to the same phonological category as the target, using the categories of velar plosive, palatal affricate, dental plosive, bilabial plosive, glide, fricative, or nasal. For example, the substitution of the voiced velar /g/ for the voiceless velar /k/ was considered a phonological error. For these and the following intra-akshara analyses, we pooled the data from the skilled and less-skilled groups. This procedure was necessary because there were few errors on some type of items.

Of the substitutions on base consonants in 60 CV akshara, 32 were classified as visual only, 115 as phonological only, 15 as both visual and phonological, and 21 as neither phonological nor visual. Thus, confusions within a phonetic class appear to account for the majority of consonant substitution errors. Among the velar plosives, palatal affricates, dental plosives, and bilabial plosives, the most common type of substitution was between voiced and unvoiced phoneme pairs, followed by confusion between aspirated– unaspirated pairs. Thus, \( \text{padaka} \) (‘medal’) was spelled as \( \text{padhaka} \) and \( \text{jalaka} \) (‘bath’) as \( \text{jhalaka} \). Aspirated sounds are not prominent in the dialect of children in the area, as mentioned earlier, and errors involving the aspirated–unaspirated distinction may reflect an influence of dialect features on spelling. In the three remaining phonological classes, most errors were between the liquids \( \text{s} \) and \( \text{l} \) (\( /\text{s}/, /\text{l}/ \)), the nasals \( \text{r} \) and \( \text{r} \) (\( /\text{r}/, /\text{r}/ \)), and the fricatives \( \text{s} \) and \( \text{s} \) (\( /\text{s}/, /\text{s}/ \)).

We observed one consonant error that appeared to be associated with another feature of the present dialect–dropping of initial /h/. The spelling list had two tokens of initial /h/, and it was omitted 31% of time. Thus, \( \text{halva} \) (‘a sweetmeat’) was sometimes spelled as \( \text{alva} \) and \( \text{hasiru} \) (‘green’) as \( \text{asiru} \). In none of the other words was the initial consonant ever dropped and substituted by a vowel. The number of cases is small, but the difference between the words with initial /h/ and the other words is striking.

Turning to vowels, we examined errors on the 17 short vowels other than the inherent vowel and the 6 long vowels, all of which are conventionally spelled with secondary or diacritic forms. On short vowels, there were 13 cases in which the diacritic form for the corresponding long vowel was substituted, as when \( * /i/ \) was spelled as \( * /i/ \). There were 30 cases in which no diacritic vowel was used, resulting in a symbol that represents the consonant with the inherent vowel. On long vowels, there were 156 cases in which the diacritic form of the corresponding short vowel was used (e.g. \( * /e/ \) spelled as \( /e/ \)) and 5 cases in which no diacritic vowel form was used. The proportion of errors involving the contrastive pair for length versus omission of the required diacritic vowel differed for long and short vowels, \( \chi^2(1) = 101.46, p < 0.001 \), with omissions more common for short vowels. These errors may arise for phonological reasons: long and short vowels are phonologically similar, and short vowels are sometimes dropped in speech. Another factor that may contribute to these errors is a difficulty in learning about and using marks that are small and not on the line of print, as the diacritic forms of
vowels often are. The diacritic forms of long vowels are elaborations of those for the corresponding short vowels, with some of them involving the addition of a separate mark. When this mark is omitted, a long vowel such as *ε turns into the corresponding short vowel, * in this example. (An analysis contrasting visual and phonological substitutions, as was done for consonants, was not suitable in the case of vowels. For vowels, substitutions that are similar on the phonological dimension are generally also similar on the visual dimension, and vice versa.)

We turn now to errors within CCV akshara, which as we have seen already is the hardest type of akshara for children to spell. The correct spelling of CCV akshara demands knowledge of two secondary forms, that for the vowel and that for the second consonant. For example, correct performance on the akshara  Zika /tre/ requires knowledge of the secondary forms of the second consonant /ʈ/ and the vowel * /e/. On the six mixed consonant clusters in our list, children performed substantially better on the vowel secondary form than the consonant secondary form, 0.67 as compared with 0.42, \(\eta(59)=6.91, p<0.001\). That is, it appears to be the clustered consonants that contribute most to the difficulty of CCV akshara.

The secondary forms of consonants are less common than the primary forms, as mentioned earlier, whereas the secondary forms of vowels are more common than the primary forms. We suspect that the low frequency of the secondary forms of consonants is one reason why it takes time for children to use them correctly in spelling.

The secondary forms of consonants differ in whether they are a miniature of the primary symbol, which is typically learned first, or whether they are entirely new symbols. We suspected that the latter type of form would be more difficult since substantial new learning is required. To test this idea, we compared performance on the three consonant secondary forms in our list that are new symbols with the five consonant secondary forms that are a miniature of the primary symbol. The proportion of correct spellings was over twice as high in the miniaturized forms (0.71) as the new forms (0.33), \(\eta(59)=15.11, p<0.001\). There are just a few examples of each type of secondary form, but the results suggest that the need to learn a new shape for a secondary form is a source of difficulty for Kannada learners. (This issue could not be examined for vowels because, as mentioned previously, the diacritic forms always differ in shape from the primary vowel forms.)

A further analysis compared performance on the eight geminates (e.g. /kk/, /bb/) and six mixed consonant clusters (e.g. /lu/, /su/) in the list. Errors on the second consonant of these clusters were coded as omissions when the consonant was dropped (e.g. Zika for /stA/ Zika, where the secondary form for /ʈ/ was omitted) and as in-line errors when the consonant was written as a primary symbol rather than a secondary form (e.g. Zika for /stA/ Zika, where /ʈ/ is symbolized with the primary form rather than the secondary form). Among geminates, omission errors greatly outnumbered in-line errors, 217 as compared with 7. Among mixed clusters, there were somewhat more in-line errors than omission errors, 85 as compared with 73. The interaction between type of error and geminate versus mixed cluster was significant, \(\chi^2(1)=101.26, p<0.001\). Among mixed clusters, in-line errors were particularly common for the two clusters that began with sonorants (/lp/, /lu/). For these clusters, there were 57 in-line errors as compared with 25 omission errors. On the four mixed clusters that began with obstruents, omission errors greatly outnumbered in-line errors (60 versus 16), as they did for geminates. The interaction between sonorant-initial versus obstruent-initial cluster and the type of error was significant, \(\chi^2(1)=35.34, p<0.001\). The numbers of clusters of different types are small, and the trends would need to be confirmed in a larger study. If confirmed, the results would suggest that some secondary forms of consonants, like those of vowels, are susceptible to omission. Clusters such as /lu/ may show a different pattern, with a number of errors in which the primary forms of both consonants are written on the line. This pattern of errors may arise because these clusters have a mismatch between the phonological syllable (e.g. hal-va) and the orthographic syllable (e.g. ha-lva). Children may sometimes (incorrectly) write the /u/ of /lu/ on the line using the primary symbol because it is the initial consonant of its phonological syllable in the same way that they (correctly) write the /u/ of /ʈu/ on the line using the primary symbol.
3.4 Summary of difficulties in learning to spell in Kannada

Ours is a first empirical study of spelling development and spelling errors in Kannada, and the results must be considered tentative. However, the findings suggest that learners of Kannada face more challenges than previously acknowledged. In this section, we summarize some of those challenges.

The children in this study, like most others in Kannada language schools, were taught akshara as whole. They were not explicitly taught how to build akshara from their components. Our results suggest that children try to do this: they are not restricted to memorization and retrieval of whole akshara. However, by the last two years of primary school, a number of children have not achieved the skills that are necessary for constructing complex akshara. If they had, they should have been more successful at spelling the akshara with vowels other than the inherent vowel and the akshara with consonant clusters. Many of these more complex akshara, especially those with consonant clusters, are less common than many of those with inherent vowels, although this impression remains to be quantified. However, if children had learned the regularities of the system and were able to use them in spelling, they should have been able to spell even uncommon akshara based on knowledge of the system. The results of Nag (2007) for reading point to a similar conclusion: complex akshara are difficult for children to read aloud even though they are formed from smaller components in a predictable manner.

Although each consonant and vowel segment has both a primary and a secondary form, the choice between the two forms is governed by rules. Our findings show that it takes some time for children to master the two forms of each segment and the rules governing their use. Children have particular difficulty when the secondary form differs from the primary form in shape as well as size, as it does for some consonants and all vowels. Also contributing to the difficulty with secondary forms is that they are often small, off the line of print, or both. Even when the secondary form is more common than the primary form, as it is for vowels, spellers sometimes omit it. Other errors reflect the use of a primary form when a secondary form is required, as in the case of certain consonant sequences.

When selecting a spelling for a phoneme, children must distinguish that phoneme from others. This is a source of difficulty for learners of Kannada, as it is for learners of other writing systems, and it points to the phonological basis of Kannada spelling. Many of the children’s substitution errors on the primary forms of consonants reflected phonological similarity, as when children used the symbol for a voiced consonant rather than the symbol for the corresponding unvoiced consonant. Confusions between phonemes that differ only in voicing have also been noted in spellers of other languages (Treiman, 1993 for English; Pollo et al., 2008 for Portuguese). The confusions that we observed between short and long vowels may reflect phonological as well as visual factors, and the representation of phoneme duration has been recognized as a spelling problem even in highly transparent (or regular) writing systems (Lehtonen and Bryant, 2005).

Mismatches between the language that a child speaks and the language that is written appear to impact the learning of Kannada, just as they impact the learning of other writing systems (e.g. Kohler et al., 2007). We found preliminary evidence for this in children’s failures to spell initial /h/, which is not pronounced in their dialect. We also found preliminary evidence that mismatches between phonological syllables and orthographic units, such as that between hal-va (phonological syllabification) and ha-lva (orthographic rendition), may be associated with spelling errors.

In summary, our data suggest that many errors at this stage of spelling development can be ascribed to phonological processing. The relatively small number of substitutions of globally similar looking akshara point to a sub-syllabic approach to spelling. Our analyses of errors at the intra-akshara level clearly demonstrate that children attempt to use phonemic information to build up akshara even though their teaching has focused on the syllable level.

3.5 Limitations, implications, and future directions

Only 30 words were examined in this study and children were from only two grade levels. The words...
were chosen to sample a number of linguistic features, but certain features were represented poorly or not at all. In future work, it will be necessary to examine specific features in more depth. It will be important, as well, to develop measures of word and akshara frequency in children’s reading materials that could be used as a basis for selecting stimuli and interpreting results.

Although much work remains to be done, our results suggest that learning to spell in Kannada is not as easy a task as some have thought it to be (Karanth, 2006). The slow pace of learning to spell and learning to read (Nag, 2007) in Kannada is surprising in light of some previous crosslinguistic studies, notably that of Seymour et al. (2003). Based on their study of reading development in thirteen European languages, Seymour et al. (2003) suggested that two factors contribute to differences in ease of acquisition across writing systems. The first is transparency of spelling–sound relations, and the second is complexity of syllable structure. Kannada is often considered a transparent writing system (Nag, 2007; Padakannaya and Mohanty, 2004). A sequence such as /ka/ or /kku/ is spelled the same way every time it occurs, and even though each phoneme has two possible spellings, the choice between them is governed by rules. Also, Kannada has a fairly simple syllable structure. Despite this, the present findings and those of Nag (2007) show that literacy acquisition in Kannada is a protracted process. Orthographic transparency, as it is often defined, does not guarantee ease of learning.

The study of Seymour et al. (2003), which examined a limited number of alphabetic writing systems, most of which used the Latin script, may have missed certain factors that contribute to the ease or difficulty with which writing systems are learned and may have oversimplified others. Orthographic transparency, we suggest, is one of the latter. It is not clear how transparency should be defined in the case of an alphasyllabary: whether at the level of phoneme sequences or at the level of individual phonemes. Currently, researchers consider a number of characteristics when describing the transparency of a writing system, including the existence of many-to-one mappings between linguistic units and spellings and the degree to which many-to-one mappings can be rationalized on the basis of morphological, phonological, and graphic context. The idea that all writing systems can be arrayed along a single continuum of transparency is questionable even for alphabets. We suggest that it is more questionable when other types of writing systems are considered.

Previous studies, including that of Seymour et al. (2003), have primarily examined literacy acquisition in systems that use the Latin script. This limitation may have caused researchers to downplay a factor that, according to our results, seems to be quite important in literacy acquisition: the visual complexity of the symbols in a writing system. Kannada and other Indian alphasyllabaries have adopted what seems to be an elegant compromise between alphabets and syllabaries, forming graphic complexes from subparts that represent individual consonants and vowels. But, given the pressure for symbols of writing to be small and quick to produce, the subparts are often quite small. Some of them are above or below the line of print, and they are often joined to one another in such a way that it can be unclear where one ends and another begins. These facts probably encourage people to treat each akshara as a unit. In that case, the large number of akshara symbols that need to be learned becomes an issue. The need to segment speech into phonological segments is a widely acknowledged stumbling block in learning to read and write, and phonemic analysis has been a major focus of research and theory with alphabetic scripts (Goswami and Bryant, 1990). Our results suggest that the need to segment the visual symbols of writing can be another challenge in learning to read and write, one that is less widely appreciated.

Perhaps because of the visual unity of the symbols in Kannada and similar alphasyllabaries, children are often taught the akshara as wholes. In many cases, they are not explicitly taught about the parts within the akshara and about how these parts signify phonemes (but see Nag and Sircar, 2008). Because of the large number of akshara in Kannada and other Indian writing systems, this teaching extends over several years. Our results suggest that, by 4th and 5th grade, children do spontaneously shift...
to an analytical approach and attempt to build up akshara from smaller components. However, they do not always have the skills that would allow them to do this accurately. Explicit instruction about how akshara are built from subparts representing consonants and vowels and about the secondary forms could benefit learners of Kannada and other Indian alphasyllabaries. Increasing children’s familiarity with the symbols through a variety of reading and writing opportunities could be helpful in this process.

The course of spelling development in Kannada may reflect characteristics of the writing system and of how it is taught, but it might also reflect home environment and school attendance. Many children in the area of our study have parents who are not literate in Kannada, and some come from homes that do not have newspapers and books. In some schools, additionally, some children’s attendance is erratic. The time taken to master some aspects of spelling in Kannada may reflect, in part, these factors.

The alphasyllabary of Kannada is not an outlier writing system. Approximately 37 million people speak the language in India alone (Census Data 2001), and a number of other writing systems in South and Southeast Asia, including Bengali, Hindi and Thai, are based on similar principles. A better understanding of literacy and its acquisition in alphasyllabic writing systems can help to counter the current Anglocentric focus of reading science (Share, 2008), providing a basis for the development of more inclusive theories.

Acknowledgements

We thank the Department of Education in Chamarajanagara District and members of the District Quality Education Project at the National Institute of Advanced Studies for their support in the field; Mohan Das, Roopa Kishen, Mamta Gupta and Kala B. from The Promise Foundation for data collection and data coding; and the members of the Centre for Reading and Language at University of York and the Reading and Language Lab at Washington University for their helpful comments on a draft of the manuscript. Finally, we thank the children who willingly participated in the study.

Funding

This work was supported by a research grant to M.J.S from the British Academy - ‘Reading Difficulties in Kannada, an Indian Alphasyllabary’.

References


Appendix A

Items for spelling test

<table>
<thead>
<tr>
<th>Kannada spelling</th>
<th>Phonemic transcription</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ಡಾಣಾ</td>
<td>dhana</td>
<td>Cattle</td>
</tr>
<tr>
<td>ಬಾಣಾ</td>
<td>bana</td>
<td>Forest</td>
</tr>
<tr>
<td>ಅಬಾಣಾ</td>
<td>abana</td>
<td>Washer man</td>
</tr>
<tr>
<td>ಹರ</td>
<td>hara</td>
<td>Heroes</td>
</tr>
<tr>
<td>ಪಾದಾಕಾ</td>
<td>padaaka</td>
<td>Medal</td>
</tr>
<tr>
<td>ಬಾಳ</td>
<td>bal</td>
<td>Bath</td>
</tr>
<tr>
<td>ಸಾವಿ</td>
<td>sawi</td>
<td>Sweetness</td>
</tr>
<tr>
<td>ಮಿ:ನಾ</td>
<td>mi:na</td>
<td>A girl’s name</td>
</tr>
<tr>
<td>ಗುರು</td>
<td>guru</td>
<td>Teacher</td>
</tr>
<tr>
<td>ಬಾಹಿಕ</td>
<td>bahika</td>
<td>Boy</td>
</tr>
<tr>
<td>ಮುದುಕಾ</td>
<td>muduka</td>
<td>Old man</td>
</tr>
<tr>
<td>ಪಾರಾದೆ</td>
<td>paraade</td>
<td>Curtain</td>
</tr>
<tr>
<td>ಬಾ:ಗಿಲು</td>
<td>bagila</td>
<td>Door</td>
</tr>
<tr>
<td>ಹಸಿರು</td>
<td>hisiru</td>
<td>Green</td>
</tr>
<tr>
<td>ಕಿರಾ</td>
<td>kirä</td>
<td>Ray of light</td>
</tr>
<tr>
<td>ಪಾಜಿಡಾ:ಮಾ</td>
<td>pajida:ma</td>
<td>Pant</td>
</tr>
<tr>
<td>ಕೆ:ರಾ</td>
<td>ke:ra</td>
<td>Name of an Indian state</td>
</tr>
<tr>
<td>ಕೊ:ಮಾಲ</td>
<td>ko:maal</td>
<td>Tender</td>
</tr>
<tr>
<td>ಹಾರಾ</td>
<td>hara</td>
<td>Color</td>
</tr>
<tr>
<td>ಮಾನಸು</td>
<td>manasu</td>
<td>Mind</td>
</tr>
<tr>
<td>ಸಕ್ಕರೆ</td>
<td>sakkare</td>
<td>Sugar</td>
</tr>
<tr>
<td>ಹಾವಾ</td>
<td>hava</td>
<td>A sweetmeat</td>
</tr>
<tr>
<td>ಕಾಪಾಂನೆ</td>
<td>kapanne</td>
<td>Imagination</td>
</tr>
<tr>
<td>ಅಸ್ಪಾಟೆ</td>
<td>aspatre</td>
<td>Hospital</td>
</tr>
<tr>
<td>ಗುಬ್ಬಾಸಿ</td>
<td>gubbasa:si</td>
<td>Sparrow</td>
</tr>
<tr>
<td>ಮೊಮ್ಮಕಲು</td>
<td>mommakalu</td>
<td>Grandchildren</td>
</tr>
<tr>
<td>ಸ್ವಾಮನ್ಟಿರಾ</td>
<td>svamantrirä</td>
<td>Freedom</td>
</tr>
<tr>
<td>ಕುಂಜಾ</td>
<td>kunja</td>
<td>Lame man</td>
</tr>
<tr>
<td>ಹೆರಾ</td>
<td>hera</td>
<td>Meaning</td>
</tr>
</tbody>
</table>