



Learning to spell a regularly spelled language is not a trivial task – patterns of errors in Kiswahili

KATHERINE J. ALCOCK¹ and DAMARIS NGOROSHO²

¹*Partnership for Child Development and Department of Psychology, City University, London, UK;* ²*MAKWAMI Project, Bagamoyo, Tanzania*

Abstract. Various theories of spelling development are discussed, including their relevance to regularly spelled languages. For those languages studied so far, models including the incorporation of a wide variety of linguistic knowledge seem most fruitful. Data from studies of reading, however, suggest that when the language is regularly spelled children do not make many errors after the initial stages. Data are presented from spelling errors in children learning to spell Kiswahili, a regularly spelled, non-European language. Patterns of errors and even specific phonemes and graphemes that are problematic are shown to resemble closely the patterns found in English and other European languages. It is concluded that, as in other languages, children are integrating many different types of linguistic knowledge in their attempt to spell words correctly; dialect, orthography, and grammatical knowledge are all important. Unlike reading such a language, spelling a regularly spelled language is a cognitively challenging task.

Key words: African languages, Kiswahili, Regular orthography, Spelling

Introduction

Learning to spell is more difficult than learning to read. It is often found that children can read words that they cannot spell, and children who are poor readers often have even more difficulties with spelling than with reading (Treiman, 1997). Spelling tasks seem to reveal more about a child's phonological knowledge than is possible with reading tasks. Languages fall on a continuum of transparency of spelling, but this has been found to hold for both fairly non-transparent languages, such as English, and for more regularly spelled languages such as Spanish or Portuguese (Bryant, Nunes & Aidinis, 1999; Mann, Tobin & Wilson, 1988; Read, 1986; Snowling, 1982; Treiman, 1993).

Models of learning to read

Many differing theories have been put forward in spelling development. Early suggestions included the idea that children must learn the order of letters in words by rote memory, thereby learning every word separately with no

role for phonological knowledge or generalisations. Errors, according to this view, are due to inadequate memory for a word, and patterns of errors, including serial position effects and length of word effects, have been taken to provide evidence for this model of spelling development. This type of model is tempting for languages such as English that have extremely deep orthographies, and where phonological knowledge or generalisation are not obviously useful in learning to spell.

More recently, alternative models have been proposed. Probably the first of these is the dual route model (Ellis, 1993). In this model, spelling, similarly to reading, is accessed by either a phonological or an orthographic route. This model is sometimes assumed not to be applicable to spelling development but there are a variety of authors who have used it, especially in the contexts of individual differences in spelling style (Castles, Holmes & Wong, 1997) and of developmental dyslexia (Masterson, 2000), but also more generally in the context of spelling development (Barry, 1994); it appears to be influential and worth examining.

Central to this model is the comparison of error rates and speed of access (the latter particularly in studies of reading) for words that are “regularly” or “irregularly” spelled. Again this type of model is tempting for deep orthographies. However since English is not wholly opaque in its spelling, especially when compared to logographic languages such as Chinese, the model includes features that take account of regular spellings as well as irregular spellings.

This type of model of spelling has also been criticised as rigid, in not taking account of the interaction between phonological skills and orthographic skills (Gill, 1992), or the way in which phonological processing prepares children for orthographic processing (Lennox & Siegel, 1994). In addition, it seems that this type of model may not be appropriate for languages that have either a much deeper or a much shallower orthography than English. However, some evidence has suggested that orthography influences spelling even in languages with highly regular orthography, and proponents of this theory have taken this as evidence for some degree of universality for the dual route model (Barry & de Bastiani, 1997).

Other theorists have suggested that learning to spell takes place in stages, with a variety of stage theories existing, including that of Frith (1980) where children pass through a logographic, a phonological, and an orthographic stage. Again these types of models have been criticised as too simplistic – children may simultaneously use different types of information, rather than discarding one type of information upon leaving the relevant stage. In addition, the evidence for some of the stages, e.g. a logographic stage of reading, has been seen as quite weak (Gill, 1992).

Linguistic analyses of spelling errors

More recently, researchers have started analysing the spelling patterns of English and children's spelling errors in a more qualitative and more strictly linguistic manner, taking into account particular phonological patterns, and this has led to new thinking about many of the previous theories. For example, many previous studies have distinguished in a bimodal way between "regular" and "irregular" words but in reality there is a continuum of words which have either the most common spelling of a particular phoneme, a less common spelling of that phoneme, or a unique spelling of that phoneme. In addition, children's spellings are often classified either as correct or incorrect, failing to take into account phonetically accurate but incorrect spellings.

Alternatively, children may have good knowledge of the set of graphemes that can represent a particular phoneme but may lack, for example, contextual knowledge. This would prevent them from realising that /ou/ can be spelled O in the middle of a word (PHONE) but not at the end of a word (LOW). Even if these factors are taken into account, and such spellings are classified as phonetic, different nonphonetic spellings are not equally likely; and different types of words are not equally likely to be spelled nonphonetically (Treiman, 1993).

Children's difficulties with spelling, even where the resultant spellings appear nonphonetic to adults, are not necessarily due to inadequate phonological knowledge. In some cases their phonological knowledge may be more accurate than that of adults – for example, they may spell "stick" as SDIK, indicating that they realise the second consonant is not the same as the initial consonant in "tick". Children's invented spellings are particularly interesting here, as children who have not been taught the orthography of a word are relying only on their own phonological knowledge (Read, 1986).

In addition, when asked to count the number of sounds in words they may be more accurate than adults – indicating, for example, that they realise that "my" has one consonant and two vowel sounds, despite being spelled with only one vowel sound; adults usually indicate that this word has two sounds (Treiman, 1991).

When children's errors are pointed out to them, some will change their spellings but many will not. A common error is spelling words with initial consonant clusters (such as "play") with only the first of the two consonants (e.g. as PAY). When this error is pointed out, and children are asked what the difference is, they will frequently say that the words sound differently, that the difference is in the region of the /p/, but that they are still spelled the same (Treiman, 1985; Treiman, 1991).

There seems to be more to learning to spell than these models can capture. It seems that children need simultaneously to integrate not only phono-

logical, but also orthographic, lexical and grammatical knowledge in order to become accomplished spellers. The same sound can be spelled in different ways depending on its grammatical function; in English children must learn that the /nd/ sound in “hand” and “tanned” has a different grammatical function and hence is spelled differently. There is evidence that more accomplished spellers use both orthographic and grammatical knowledge more than beginning spellers (Nunes, Bryant & Bindman, 1997c).

However, not all languages are as difficult to spell as English and it is possible that these many sources of information are not as important in regularly spelled languages. The use of orthographic information in spelling Italian has been discussed (Barry & de Bastiani, 1997). Grammatical knowledge can also be shown to be important in some regularly spelled languages, such as Portuguese (Pinheiro, 1995). Knowledge about orthography, for example which double letters are permissible, is usually found only in more experienced spellers, with beginning spellers in English willing to accept HH as a permissible letter combination. However, even beginning spellers have some orthographic knowledge, and will often indicate words that start with double letters to be unacceptable spellings. (Bryant et al., 1999; Cassar & Treiman, 1997; Rego & Bryant, 1993).

One particularly interesting set of errors found in beginning spellers is the tendency to represent consonant clusters by using only one of the consonants. In English word-initial or syllable-initial clusters are usually stop-liquid (as in “play”) or fricative-other consonant (as in “smell”, “fly” or “step”). In syllable-final clusters the liquid or nasal often appears first, before the stop or fricative (as in “paunch” or “curt”) although again clusters with a fricative followed by a stop are found (as in “past”). In word- or syllable-initial consonant clusters in English, children tend to omit the second of the consonants. However, when word- or syllable-final clusters that start with nasals or liquids are examined, it is the first of the two consonants that are usually omitted. Hence children are likely to spell “star” as SAR but spell “sand” as SAD (Treiman, 1993).

Beginning literacy in other languages

Languages differ not only orthographically but also in terms of their overall phonological patterns – not only which phonemes exist in the language, but which combinations of these phonemes are permissible, and what constitutes an acceptable syllable, among other differences. Hence the above clusters that are difficult for beginning English spellers may not exist in other languages. This type of difference may well prove extremely relevant to the study of learning to spell (Treiman, 1993). There may be differences in the rate and developmental trajectory of learning to spell, and there may also be differ-

ences in the types of errors found, depending not only on orthography but also on the phonological patterns that are legal in a language.

Studies of beginning readers, however, appear to show that the regularity of a language influences the ease of learning to read, with children who are learning a wide variety of regularly spelled languages being able to read all or almost all words once they have “cracked the code”, and the cluster and syllable patterning of a language being less important (Alcock et al., 2000; Oeney & Durgunoglu, 1997; Thorstad, 1991). From previous studies (Bryant et al., 1999), it seems that this pattern may not emerge for beginning spellers in regularly spelled languages – the difficulties they encounter seem to be similar to beginning spellers of more difficult languages such as English.

Studies in English have therefore found errors that can be accounted for on the basis of lack of orthographic knowledge, differing phonological representations, dialect differences, lack of knowledge about how grammatical differences are represented in the orthography, and lack of word-specific knowledge. English is of course one of the least regularly spelled languages and any conclusions about spelling development must take this into account. It is particularly interesting, therefore, to investigate regularly spelled languages – especially unrelated, non-European languages – to determine whether these linguistic facts must be learned by beginning spellers in such languages in order to be successful spellers. Similar studies have been carried out in European languages that are regularly spelled (at least from the point of view of reading – as with Kiswahili, the language under investigation in this research, some of them present spelling difficulties; (Harris & Giannouli, 1999; Justicia, Defior, Pelegrina & Martos, 1999; Pinheiro, 1995) but an extension beyond this narrow grouping of the world’s languages also seems profitable. The current study investigates patterns of spelling errors in Kiswahili.

Education in Tanzania

A brief explanation of the Tanzanian education system may be helpful in explaining the study design. Primary schooling consists of seven years of theoretically compulsory and theoretically free education, starting nominally at the age of seven years. However, although over 85% of children receive some primary schooling, only 65% of children of the correct ages (7–13 years) are in school, and only 10% of children are in the correct age for their grade (The Partnership for Child Development, 1999). Most children enrol late, with the modal number of years behind correct age for grade being three, i.e. the modal age at enrolment is 10 years, but there is great variability. The reasons for late enrolment are various but include family finances – there are various payments associated with schooling as well as loss of income from

a potentially economically active member of the family – and distance from school, where there is not a school in every village (Alcock, 2001).

Many children in this population do not know their own age nor their date of birth, and in some cases parents do not know the date of birth accurately. While every effort was made to find the children's dates of birth, in almost all cases this could not be determined more accurately than one calendar year.

However, previous research (Ceci, 1991) suggests that in many studies that investigate developmental effects by examining changes with age in the school population, the studies are in fact investigating the effect of years of schooling, but these cannot be disentangled in populations where years of schooling correlates perfectly with age. In populations where these two variables are not in one-to-one correspondence, years of schooling is found to have a greater impact than years of age. Hence in this study it was decided to examine the effect of years of schooling as well as years of age, since the number of years of schooling may have a greater effect on skills than the number of years of age, and this does not correlate perfectly with years of age in the way that would be expected in a Western population.

Kiswahili – a short background to the language

Kiswahili is a language of the Bantu group spoken in areas of East Africa, particularly Tanzania and Kenya. It is used as the medium of primary education in most regions of Tanzania, is taught in schools in Kenya and Uganda and is a trade language in a wider area of East Africa, including some areas of Zaire, Burundi, Rwanda, Mozambique and Zambia. However, use as a first language is limited to coastal and island areas of Tanzania and Kenya, with more recent use in conjunction with other local languages in major conurbations in Tanzania.

Kiswahili, like other Bantu languages, has an extensive system of prefixes, suffixes and infixes, particularly found on verbs and to some extent derived nouns and adjectives. Most of these affixes are grammatically obligatory although some can be optional in some circumstances, in some cases depending on stylistic variation. Kiswahili also has a richly inflected noun class system. Prefixes and infixes must agree with the relevant noun, whether they are on adjectives, verbs, or pronouns, and there are between 10 and 20 noun classes, counting singulars and plurals separately as their agreements are different – it is difficult to give a precise figure as some classes have the same phonetic realisation though they are historically different, and this historical difference can lead to some differences, for example in plural formation or agreement.

As with many other Bantu languages, Kiswahili has only five vowels, and it is therefore possible to write the language using the Roman alphabet

with one-to-one grapheme-phoneme correspondence. In the early part of this century, a colonial committee transcribed the language with the intention of having regular spelling (Inter-territorial Language (Swahili) Committee to the East African Dependencies, 1956). For the most part this has been achieved, and it is possible to read the language using one-to-one simple grapheme-phoneme correspondences, with a few digraphs, but without any additional orthographic knowledge. Appendix 1 shows the phonetic realisation of each grapheme; letters do not have names as such but some classroom beginning reading instruction involves breaking suitable words down into CV syllables and the children are taught letters in the context of the pronunciation of each syllable, or in the case of vowels of each vowel, in a similar manner to the teaching of letter names in other educational systems. Many classroom teachers start with, or teach exclusively, the CV syllables CA so that for example the letter B will be first encountered in the context of the syllable /ba/, written BA.

Generally the phonological realisation of a grammatical morpheme and its graphemic realisation correspond. For example, the prefix /m/ places a noun or adjective in class 1, which contains primarily animate nouns including people and animals. Before a vowel its phonetic realisation changes to /mw/. This is reflected in the spelling, as follows:

<i>mnene</i>	“fat person”
<i>mwembamba</i>	“thin person”

However there are one or two ambiguities in spelling which must be resolved with recourse to grammatical knowledge. For example, the relative pronoun is realised as an infix which varies according to the noun class of the referent. Some variations are transparent phonologically but some are not, even though there are differences in the spelling, as follows:

a) “The book which he saw”

Kitabu alichooona

Book	he	past	which	see
<i>Kitabu</i>	<i>a-</i>	<i>li-</i>	<i>cho-</i>	<i>ona</i>
Book	3rd person	past tense	relative pronoun	see
Class 7	singular subject	marker	Class 7	
	prefix			

b) “The door which he saw”

Mlango alioona

Door	he	past	which	see
<i>Mlango</i>	<i>a-</i>	<i>li-</i>	<i>o-</i>	<i>ona</i>
Door	3rd person	past tense	relative pronoun	see
Class 3	singular subject	marker	Class 3	
	prefix			

c) “The bicycle which he saw”

Baisikeli aliyoona

Bicycle	he	past	which	see
<i>Baisikeli</i>	<i>a-</i>	<i>li-</i>	<i>yo-</i>	<i>ona</i>
Bicycle	3rd person	past tense	relative pronoun	see
Class 9	singular subject	marker	Class 3	
	prefix			

The verb in a) is distinct from the verbs in b) and c) both phonemically and in its graphemic realisation. However, because of the /l/ preceding the relative pronoun in examples b) and c), there is automatically a glide between /l/ and /o/. This means that the verbs in b) and c) are phonologically identical, and spellers must rely on morphology to determine the orthography, O or YO respectively.

The glide is also present in many other vowel-vowel transitions (there are no diphthongs in Kiswahili, and two adjacent identical vowels are also articulated separately; hence *kaa* has two syllables), but is only written in some, grammatically prescribed, circumstances, such as that above. Children can learn to use grammatical information to determine spelling even where there are two possible graphemic representations of one phonological entity. There is evidence from other languages, including regularly spelled languages, that beginning spellers do not use such grammatical knowledge but more advanced spellers can begin to do so (Bryant, Denne, Ledward & Nunes, 1997).

Historical changes and dialect differences mean that there are also a few other cases where more than one grapheme can correspond to one phoneme, and lexical or morphological knowledge is necessary for correct spelling. For example, in the dialect spoken in the study area, /h/ tends to be dropped. In order to learn which words ought to be spelled with an H, children can rely on two sources of information:

- 1) Lexical information. They can learn which common words are spelled with an H and which have an initial vowel, even if the /h/ is not pronounced in their dialect.
- 2) Morphological information. Initial /h/ tends to occur in negative verbs and verbs of habitual action, and children can use this information to determine if they should write a word with an H.

In this dialect, other phonological changes that mean learning to spell is a more difficult task are as follows: /θ/ tends to be pronounced as /s/ and /ð/ tends to be pronounced as /z/; /l/ and /r/ tend to be confused. The latter is probably a result partly of the late acquisition of this contrast in Kiswahili as in most languages (at around four and half years in English, where children are likely to have good exemplars in their input; Grunwell, 1982), and also the exposure of children to the speech of adults whose first language has only one of these phonemes. Such languages are common in other areas of Tanzania, and many adults with these languages as their first language have settled in the study area. Children even with Kiswahili as their only language are therefore likely to hear conflicting exemplars of some words, which will add to their problem representing this already difficult contrast.

There are some grammatical morphemes that are spelled with L; however no grammatical morphemes are ever spelled with R. Hence it is possible to use grammatical knowledge to decide in some cases if L or R is the appropriate spelling. In other cases, and in all cases of the use of /θ/ (spelled correctly with TH) and /ð/ (spelled correctly with DH), spellers must rely on lexical knowledge to determine the correct spelling.

Literacy in Kiswahili

Previous research (Alcock et al., 2000) has shown that children learning to read Kiswahili display the typical pattern for children learning to read a regularly spelled language, in that they can decode all words – including those they do not comprehend – once they have grasped the grapheme-phoneme correspondences. However, as in other regularly spelled languages, it is possible that spelling will not display this all-or-nothing characteristic. In addition to investigating the unique features of the language described above, it is also interesting to investigate whether errors are made at all by children who have some spelling skill, and to categorise the types of errors made, the contexts in which these occur, and the way in which children can use their linguistic knowledge to aid them in their spellings.

Predictions

Hence in this study it was decided first to investigate the types and frequency of spelling errors made by beginning Kiswahili spellers. Initially the effects of differing linguistic influences on spelling would be investigated in a broad sense, to determine if children learning to spell Kiswahili are making errors of the same types, and for the same reasons, as children learning to spell English. Included in this investigation would be an examination of phonological, orthographic, grammatical and dialectic factors in children's spelling errors.

Some brief details of the errors that are predicted are as follows. Based on dialect features, children will miss out H, confuse L and R, and spell TH as S and DH as Z. This corresponds to the findings in English on dialect influences on spelling (Treiman, Goswami, Tincoff & Leever, 1997).

Based on orthographic features, children are predicted to insert and delete Y and W between syllables and between vowels more frequently than in other positions. Based on phonological features and common types of errors in other languages, children will omit one letter of consonant clusters, and as all consonant clusters consist of a nasal consonant followed by another consonant, these omissions will preferentially consist of nasal omissions. This is analogous to children's omissions of consonants in clusters in English (Treiman, Zukowski & Richmond-Welty, 1995). Based on grammatical features, it is predicted that these types of errors will be mitigated by grammatical knowledge, such that L and R will be more accurately represented in grammatical morphemes, and Y and W more accurately represented at morpheme boundaries. It is also predicted that while all of the above will differentially improve with increasing age and/or grade level or spelling skill, those based on grammatical knowledge will preferentially improve; both of these predictions correspond to previous findings in both English and Portuguese (Nunes, Bryant & Bindman, 1997b; Bryant et al., 1999).

In addition, children's spellings in two different tasks (free writing and a structured spelling test) would be collected, in order to investigate errors in these two situations that place rather different demands on children (Snowling, 1994). In the structured spelling task, a representative sample of different types of words representing common errors found in free writing was included, for comparison between types of words, and with words that have no commonly misspelled patterns.

Experiment 1

Subjects

In total 41 children took part in this experiment, from second through sixth grades at one of two rural primary schools in Bagamoyo, Tanzania. In common with all the other children in the study, they spoke Kiswahili as their first or primary language, and were currently attending Kiswahili medium primary school. The ages of children for whom a date of birth was available ranged from 9 years to 18 years (mean 13.65 years), and these children were all in their second through fifth year of primary education (no date of birth was obtainable from any child in grade six). The age ranges in each grade were as follows:

Grade	Minimum age	Maximum age
2	9	14
3	11	13
4	12	18
5	13	15

As discussed in the Introduction, this wide variation of ages within grades is common in Tanzanian schools.

Methods

Children were asked by an educational tester to write a suitable message and draw a picture as a farewell to the first author who was leaving the project. The children were given paper and pencils, and wrote continuous prose, drew and labelled their drawings. Children's work was not checked by teachers or research assistants, and they were not specifically told to check their spelling. However owing to the transparent nature of the orthography of Kiswahili, the reasonably predictable intent of the children's messages, all written on the same theme and many in a slightly formulaic style, the drawing by some children of labelled pictures, and the moderately low number of errors per word, it was relatively easy to determine the children's intended meaning.

Words written were then scored as correctly spelled or not, with any errors categorised by what morpheme they occurred in (whether they occurred in a root morpheme, a bound morpheme or at a morpheme boundary) and what type of error occurred – what letter was spelled incorrectly, and whether the error was one of omission, addition, or substitution.

Table 1. Experiment 1, spelling errors by grade.

Grade	Total words attempted	Words correct	Percent correct	Different words attempted	Type token ratio
Not given*	66	54	81.8%	47	0.71
2	38	28	73.7%	24	0.63
3	77	54	70.1%	44	0.57
4	240	196	81.7%	91	0.38
5	277	239	86.3%	74	0.27
6	91	81	89.0%	66	0.73
All children	789	652	82.6%	196	0.25

*Some children did not write their grade on their drawing/writing sample; some did not even write their name, and hence could not be checked in school records, for some with a name there was more than one child in school records with the same name.

Results

Overall results and trends by grade

In total the 41 children attempted to spell 789 words, an average of 19.24 words each. Of these words, 652 or 82.6% were spelled correctly. Table 1 shows the breakdown by grade.

In higher grades, there was a non-significant trend for children to write more words and, although they did write significantly more correct words (Kruskal Wallis Chi-square = 10.562, d.f. = 4, $p = 0.032$), because they wrote more words overall the proportion of words written correctly did not change with grade. There were also non-significant trends for children in higher grades to write words that were longer in terms of syllables and letters, and there was a significant trend for children in higher grades to attempt words that had more morphemes (Kruskal Wallis Chi-square = 13.336, d.f. = 4, $p = 0.010$).

Table 2 shows total words spelled and success by age – where there were fewer than 5 children in an age group they have been grouped together.

Again there were nonsignificant trends for children who were older to attempt more words and get more words correct, and the trends for older children to write words that were longer in letters and in syllables approached significance (Kruskal Wallis Chi-square = 10.437, d.f. = 5, $p = 0.64$ and Kruskal Wallis Chi-square = 10.760, d.f. = 5, $p = 0.56$ respectively), although the trend for older children to write words that had more morphemes was not significant.

Table 2. Experiment 1 – Spelling errors by age group.

Age	Mean words attempted	Mean words correct	Percent correct
9–11	12.00	8.67	75.9%
12	21.00	15.00	71.7%
13	11.86	10.57	88.0%
14	29.50	24.40	81.7%
15	16.00	14.20	85.7%
16–18	5.00	4.00	80.0%
All children for whom age known	19.03	15.56	0.815

Types of words

Children attempted personal names and place names as well as other words. Treiman (1993) found that proper names were easier to spell than other types of words. Here we found that personal names were spelled correctly less often than other types of words (Mann-Whitney $U = 8667$, $p = 0.003$); however place names were spelled correctly more often than other types of words (Mann-Whitney $U = 26131$, $p = 0.026$). It was possible that influence from English personal names might have contributed to the lower accuracy with personal names, but none of the children attempted to spell an English derived personal name; there were also no grade differences in the accuracy of either personal names or place names.

Error types

Examples of all error types described here can be seen in Table 3. The most common type of errors was a confusion between L and R, accounting for 66 errors out of 137, or 48.2%. Addition of Y or W accounted for a further 16 errors, or 11.7%, and omission of H 11 errors, or 8.0%. The other common types of errors included omission of Y or W and addition of H. These can be seen in Figure 1.

Error place

The most common part of the word for errors to occur in was the root morpheme. However the type of error significantly interacted with the place of the error, with L or R substitutions being most likely to occur in root

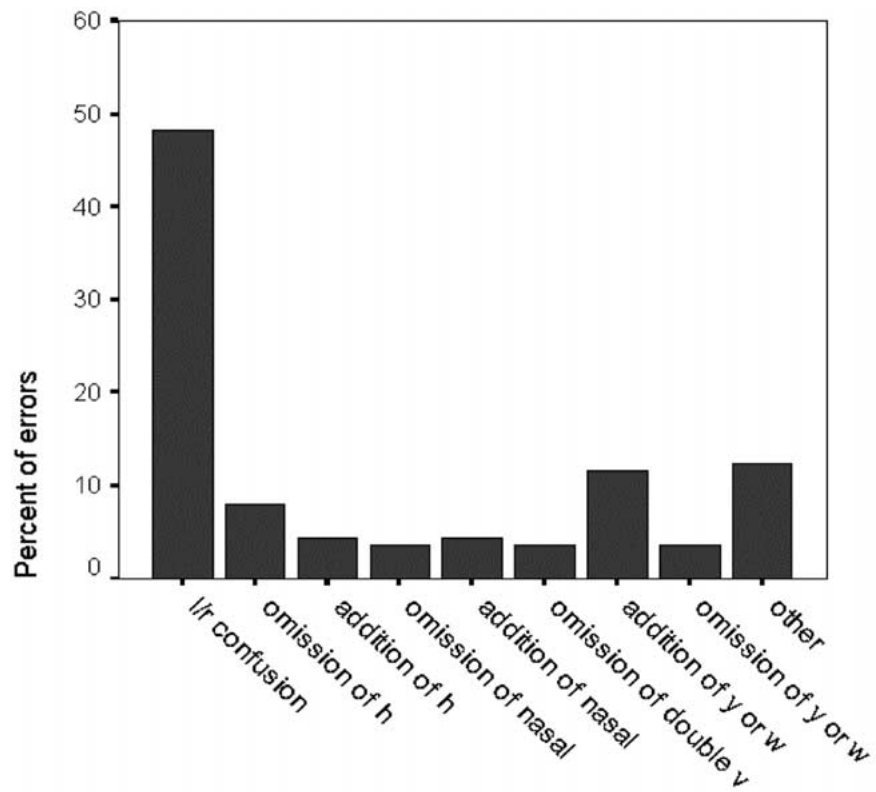


Figure 1. Distribution of error types in experiment 1.

Table 3. Experiment 1 – examples of error types.

Error type	Swahili word	Phonetic form	Spelled as	English meaning
L/R substitution	<i>heri</i>	/herɪ/	HELI	“happiness”
Y or W addition	<i>ua</i>	/uwa/	UWA	“flower”
H omission	<i>hapa</i>	/hapa/	APA	“here”
Y or W omission	<i>akupendaye</i>	/akupɛndajɛ/	AKUPENDAE	“one who loves you”
H addition	<i>uende</i>	/uwendɛ/	HUENDE	“you should go”

morphemes, H omissions being most likely to occur in bound morphemes and Y or W additions being most likely to occur at morpheme boundaries (K-W Chi-square = 25.972, d.f. = 2, $p < 0.001$).

Developmental change in error types and places

For the more common errors, differences between grades for error frequency were analysed. The addition of H did not change with grade, nor did L/R substitution errors. Addition of Y or W and omission of Y or W both altered with increasing grade (chi-square = 30.956, d.f. = 4, $p < 0.001$ for addition and chi-square = 15.985, d.f. = 4, $p = 0.003$ for omission). However in both cases there was a slight increase in this type of error with increasing grade.

For the errors involving Y and W, it was found that children in older grades were more likely to make these errors increasingly in bound morphemes rather than in root morphemes (for addition of Y or W, chi-square = 7.969, d.f. = 4, $p = 0.047$; however for omission of Y or W all errors were made in bound morphemes).

Length of word

Longer words were also harder to spell, whether length was measured in letters (d.f. = 12,776, $F = 7.691$, $p < 0.001$), syllables (d.f. = 6,782, $F = 12.645$, $p < 0.001$) or morphemes (d.f. = 6,782, $F = 8.350$, $p < 0.001$).

Less common errors

A number of error types were found with a frequency greater than zero, but which were not very common. These can be seen in Table 4. One infrequent error type is worth mentioning further. Treiman (1993) found that words with digraphs (two letters that spell a single phoneme) were more difficult to spell. This was not found here; there was no significant difference between the level of success at spelling words with and without digraphs. However, one particular type of digraph (TH or DH) is possibly more difficult to spell, based on dialect features. There was only one attempt at a word containing this type of digraph, which was a proper name, and which was misspelled both times it was attempted, even though it was the child's own name.

Table 4. Experiment 1 – Frequency of less frequent error types.

Error type	Frequency (expressed as % of opportunities to make error)
Nasal consonant omitted	1.5%
Nasal consonant added	0.8%
Double vowel spelled as single	16.7%

Experiment 2

Subjects

A total of 188 children took part in this experiment, all in second and third grade in five different schools. The children ranged in age from 9 years to 15 years, with a mean age of 12 years, age range in grade 2 being 9 to 13 years and in grade 3 being 9 to 15 years. This is a normal age range for these grades in rural Tanzanian primary schools. Children were tested in April or May. The school year runs from January to December and hence the children had either one year and four months or two years and four months of schooling.

Methods

This task was adapted from the WRAT (Jastak & Jastak, 1965; Jastak & Wilkinson, 1984) spelling task. This task consists of a set of 50 words which are dictated, including a context sentence for each one. The words were arranged in increasing order of difficulty following piloting, and a sample of the different orthographic combinations commonly found to be difficult by beginning spellers were included. These were included in proportion to their frequency in the language, as shown in Table 5 (below).

The tester explained to the children that they would hear a word, followed by a sentence containing the word, and then they would hear the word again. They were told that they should listen carefully and only write the single word they heard at the end. An example was then given with the tester reading a word, its context sentence and the word again, and then writing the single word on the blackboard. The tester also checked the children's test papers while they were writing to ensure they were writing the required word and no others.

The tester read the sentences to the children, who were tested in groups of 30 or less. Testers were Tanzanian primary school teachers, trained in educational testing, who had Kiswahili as their first or primary language.

After scoring, children who had made very few legible attempts at any word were excluded (these children are not included in the above total), and error types were coded.

Results

Overall error rate

A total of 188 children spelled 50 words each, making a total of 9,400 words. The mean number of words correct was 40.77 out of 50.

Error types

Examples of the most common error types can be seen in Table 5.

A total of 50 different types of errors were made by these 188 children on the 50 item spelling test. These errors were ranked according to frequency firstly by the total number of times the error was made and secondly, where possible, by the proportion of words in which an error was made out of those words in which the error could possibly occur. For example, the number of words in which /t/ was spelled as L was divided by the total number of words which contained an /t/. However, for errors involving adding letters, this method of estimating error frequency is obviously not possible, and only the total number of occurrences of the error is quoted.

The 20 most frequent error types are summarised in Table 5. Where it is possible to express the error as a proportion of words in which the error can occur, this is also shown. Only those error types which occur in the most frequent 20 types both by total errors and by proportion are shown, apart from errors involving adding a letter. Errors are shown in decreasing order of overall frequency.

Less frequent errors

The frequencies of less frequent errors can be seen in Appendix 2.

Changes with grade level

There were no significant effects of grade level on the overall number of errors made by children. However, since there were only two grades represented it is possible that some developmental progression is being masked by the low variability in grade level. For further analyses of types of skill improvement children were divided equally into two groups of better and worse spellers, those with an overall score of less than 40, or 40 and over.

Table 5. Experiment 2 – Examples and frequencies of error types.

Error type	Example	Number of words out of 50*	Phonetic form	Spelled as	English meaning	Mean no of errors per child	Total errors in data set	Mean proportion of errors per child (of words of this type)
DH (/ð/) spelled as Z	<i>dhamira</i>	1	/ðæmIɾæ/	ZAMIRA	“purpose”	0.61	114	0.61
TH (/θ/) spelled as S	<i>thamani</i>	1	/θæmənt/	SAMANI	“value”	0.56	106	0.56
L (/l/) spelled as R	<i>lazima</i>	6	/ləzImæ/	RAZIMA	“necessary”	0.95	179	0.16
R (/r/) spelled as	<i>kufikirika</i>	5	/kufIkIrIkæ/	KUFIKILIKA	“to think”	1.72	324	0.34
Y omitted	<i>yasiyovumilika</i>	3	/jæsIjovumIlIkæ/	YASIOVUMILIKA	“intolerable”	0.29	55	0.07
H deleted	<i>muhimu</i>	6	/muhImu/	MUJMU	“important”	1.41	266	0.20
Nasal consonant substitution	<i>mganga</i>	34	/mɔŋæŋgæ/	NGANGA	“traditional doctor”	0.72	136	0.02
Nasal consonant deleted	<i>ndani</i>	34	/ndænl/	DANI	“inside”	0.58	109	0.01
Double vowel spelled as single	<i>kaa</i>	5	/kææ/ [bisyllabic, stress on first /æ/]	KA	“sit”	0.48	90	0.06
W added	<i>huenda</i>		/huendæ/	HUWENDA	“goes”			

Table 5. Continued.

Error type	Example	Number of words out of 50*	Phonetic form	Spelled as	English meaning	Mean no of errors per child	Total errors in data set	Mean proportion of errors per child (of words of this type)
Single vowel substitution	<i>jikoni</i>	All 50 words	/dʒɪkɔnɪ/	JEKONI	"in the kitchen"	1.11	209	0.01
H added	<i>ukingo</i>		/ukɪŋɡo/	HUKINGO	"edge"	0.37	69	
Consonant manner of articulation changed	<i>muhimu</i>	All 50 words	/muhɪmu/	MUKIMU	"important"	0.35	66	0.01
More than one articulation change to consonant (e.g. manner plus voicing)	<i>eleza</i>	All 50 words	/elezə/	ELEYA	"explain"	0.30	56	0.01
Vowel added between consonants	<i>treni</i>	17	/tɾɛnɪ/	TIRENI	"train"	0.25	47	0.01
Y added	<i>tabia</i>		/tɛbɪjə/	TABIYA	"character"	0.19	36	

*Some words contained more than one difficult spelling pattern.

There were 42 poor spellers and 53 good spellers in grade 2, and 33 poor spellers and 60 good spellers in grade 3. Using a chi-square test the proportion of good to poor spellers was not found to be significantly higher in grade 3.

Age and spelling skill effects on common error types

Words with digraphs

Words which have digraphs were more difficult to spell correctly than those which have no digraph (Wilcoxon $Z = 10.619$, d.f. = 187, $p < 0.001$). There were particular difficulties with TH and DH. These were the most common type of error, as a proportion of the number of words containing these combinations of letters. Over half the children spelled TH as S (56.3%) and DH as Z (60.6%). Children in standard 3 were no less likely to spell these phonemes incorrectly. However, when the children were divided into good spellers and poor spellers, good spellers made significantly fewer of both types of errors (Mann-Whitney $U = 3418.50$, $p = .009$ for TH and Mann-Whitney $U = 2590.50$, $p < 0.001$ for DH). However, there were also some errors on other types of digraphs (SH and CH). Five words in the test had digraphs – one each with TH and DH – and the mean number correct out of these five words was 2.91.

Errors involving L and R

These were the next most common type of error, with 34.5% of attempts to spell R resulting in L, and 15.9% of attempts to spell L resulting in R. Children were more successful in spelling L when it was part of a bound morpheme than when it was part of a root morpheme. (Wilcoxon $Z = -3.7096$, d.f. = 187, $p = 0.002$).

An extra year of school did not improve the spelling of these letters, but good spellers in both grades were significantly less likely than poor spellers to exchange R for L (Mann-Whitney $U = 2350.50$, d.f. = 1, $p < 0.001$); and there was a trend for children who were better spellers to make even fewer errors involving L in bound morphemes (Mann-Whitney $U = 3603.50$, d.f. = 1, $p = 0.052$).

Errors involving H

Here 20.2% of attempts to spell H resulted in omission of the letter, and 7.3% of words overall had an H added. Again, better spellers were significantly less likely to drop H (Mann-Whitney $U = 1098.00$, d.f. = 1, $p < 0.001$) but there was no difference between the good and poor spellers for adding H. There was however an improvement with grade level for H addition (Mann-Whitney $U = 3785.00$, d.f. = 1, $p = .017$).

Errors involving Y or W

Here 7.3% of attempts to spell a Y resulted in omission and 3.2% of attempts to spell a W resulted in omission. W was more likely to be added at a morpheme boundary than in a root morpheme (Friedman Chi-square = 99.656, d.f. = 2, $p < 0.001$). Y was more likely to be added in the root morpheme or at the morpheme boundary (Friedman Chi-square = 7.589, d.f. = 2, $p = 0.022$). Grammatical knowledge is needed in order to determine whether some words should have Y or W at the morpheme boundary, whereas placement of Y or W at other points in the word does not depend on grammatical knowledge but simply semantic or orthographic information about a particular word.

Increased spelling skill reduced the probability of adding W (Mann-Whitney U = 3062.00, d.f. = 1, $p < 0.001$), and further analysis revealed that this applied both to bound morphemes (Mann-Whitney U = 4068.00, d.f. = 1, $p = 0.033$) and to the morpheme boundary (Mann-Whitney U = 3360.00, d.f. = 1, $p = 0.005$)

Errors involving nasals

Words that contained nasals included those with nonsyllabic nasals as the first consonant of clusters, both word-medial and word-initial, those with syllabic nasals forming the first consonant of a cluster, in this case all word-initial, and single nasals that were not in clusters. The most common type of nasal error was deletion of the nasal consonant.

Treiman (1993) suggests that the probability of consonant errors occurring depends on their position in consonant clusters. The proportion of errors involving nasals, and, where appropriate, the second, non-nasal, consonant in the cluster, was calculated. These can be seen in Figure 2.

Nasals in Kiswahili can be either syllabic or nonsyllabic. Nasals in clusters, whether syllabic or nonsyllabic, were significantly more likely to be spelled incorrectly than single nasals (not in clusters) (Wilcoxon Z = 3.912, $p < 0.001$ for nonsyllabic nasals, Wilcoxon Z = 3.041, $p = 0.002$ for syllabic nasals). In addition, the nasals in clusters were significantly more likely to be spelled incorrectly than the other (second) consonants in these clusters (Wilcoxon Z = 4.075, $p < 0.001$ for nonsyllabic nasals, Wilcoxon Z = 2.657, $p = 0.008$ for syllabic nasals).

More errors were made on the nonsyllabic nasals than the syllabic nasals (Wilcoxon Z = 4.675, $p < 0.001$). In consonant clusters containing a nasal and another consonant, the other consonant was more likely to be spelled incorrectly than were non-nasal consonants in general (for nonsyllabic nasals, Wilcoxon Z = 2.539, $p = 0.011$, for syllabic nasals Wilcoxon Z = 6.468, $p < 0.001$). However there was no difference between the error rates on second consonants in syllabic versus nonsyllabic nasal clusters.

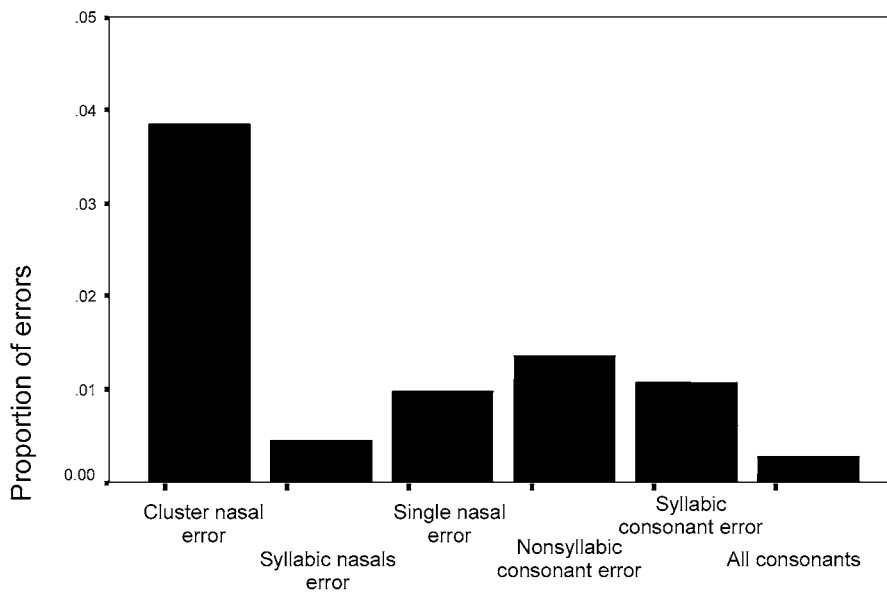


Figure 2. Errors in nasals and other consonants in experiment 2.

Hence nasals were more difficult to spell when they were in clusters than alone, whether they were syllabic or nonsyllabic; however the nonsyllabic nasals were even more difficult to spell. The second consonant in the cluster was more difficult to spell than similar consonants not in clusters, but there was no difference in the error levels in the second consonant in the two different types of clusters. In both types of nasal clusters more errors were made on nasals than on the second consonants in those clusters.

Discussion

In this study of a non-European language, which has a number of features that are not found in previously studied European languages, children learning to spell seem to make errors for the same reasons as those found in children learning to spell both regularly spelled and inconsistently spelled European languages. In many cases errors are made that are phonologically equivalent to errors made in English and other European languages, and the types of knowledge needed for good spelling seem to be equivalent in the two sets of languages.

Summary of results

In experiment 1 the most common errors were confusion of L and R, addition of Y or W, and omission of H, followed by omission of Y or W and addition of H.

In experiment 2, the most common errors were those involving the digraphs TH and DH, followed by L and R. Spelling of L was easier when it was in a grammatical morpheme – a context in which only L can occur. Omission and addition of H were also common, as were omission of Y and W, as in experiment 1.

A further type of error in experiment 2 was the omission of a nasal consonant in a consonant cluster, especially where the nasal consonant was not syllabic.

Spelling is therefore different to reading in regularly spelled languages, in that phonological code-breaking is not the only skill needed for successful spelling, while it is a sufficient skill for successful reading (Cossu, Gugliotta & Marshall, 1995). Dual route models have been proposed for skilled spelling in regularly spelled languages (Barry & de Bastiani, 1997), as well as developmental models involving stages of spelling. It appears that multiple facets of language knowledge are simultaneously necessary to achieve good spelling in even this regularly spelled language with a relatively simple syllabic structure.

Developmental trends

Overall features of the two experiments included some trends for children to spell more words correctly following more years of education, however although this trend was significant in Experiment 1, it was not significant in Experiment 2, and the proportion of words spelled correctly was not significantly greater in higher grades in Experiment 1 because children also seemed to be more adventurous in the words they chose to spell.

There were however differential effects on the type of errors made following increasing spelling skill in experiment 2, where the range of grade levels was small but the difference in spelling skill between good and poor spellers appeared to be reflecting selective improvements in particular types of skill. It appears that development in spelling skill is not automatic with increasing age or following increasing exposure to instruction, but that development in spelling skill reflects particular types of knowledge about language. This finding accords with other data from English, where increasing spelling skill overall is mainly due to increases in particular types of linguistic knowledge (Bryant & Kemp, 2003).

Comparisons with other languages

In both experiments children were found to make some errors based on dialect features, substituting L for R and vice versa, and dropping H. This effect has been observed in English (Treiman et al., 1997), and in both English and Kiswahili there is also overcompensation for the effect of dialect; this is manifested by adding H where this is not needed, in Kiswahili (e.g. *amri* “order” becomes HAMRI). Likewise in English children whose dialect omits /r/ in some contexts these children add R in words that do not have /r/, for example spelling “china” as CHINAR.

Some of the errors found in Kiswahili seem to be based on a possibly incomplete (or, alternatively, non-adult) knowledge of the language’s phonology. For example, nasal consonants were frequently omitted in consonant clusters, and there were some omissions of the second consonant in such clusters. Again, these are similar to errors observed in English (Treiman et al., 1995).

However, children are able to use both orthographic and grammatical knowledge to mitigate the effects of these errors. For example, children appear to know that grammatical morphemes cannot contain R; even though they confuse /l/ and /r/ in their spellings, they are more likely to spell grammatical morphemes with L than with R.

The letters Y and W are frequently inserted where two vowels are adjacent. This can only occur between syllables, but can be either within a root morpheme or at a morpheme boundary. Between some pairs of vowels there is a glide in spoken Kiswahili, but this is only represented in some contexts. For example, it is never represented between the last two vowels of a verb, and is represented as part of a relative pronoun only in some grammatical contexts.

This type of error was less common within a bound morpheme, as these are usually monosyllabic. However there are contexts where Y or W can legally occur between two vowels in some words, and these are at morpheme boundaries. Again, children can be seen to be using their linguistic knowledge, as we see they are more likely to place these extraneous Y and W at morpheme boundaries than at vowel transitions elsewhere in the word. In addition, children’s graphemic representations of words with these vowel transitions are phonologically more accurate than adults; all of the phonological representations contain /y/ or /w/ but adults only represent some of these.

In experiment 2 increased spelling skill differentially reduced the number of errors involving addition of W in bound morphemes or at a morpheme boundary, suggesting that these better spellers are again using their grammatical knowledge more effectively to improve their spelling.

Again, similar errors and trends have been observed in children learning to spell English (Nunes, Bryant & Bindman, 1997a) and Portuguese (Bryant et al., 1999).

Children were less accurate at spelling personal names than other types of words, in contrast to previous research in English (Treiman, 1993), although they were more successful at spelling place names than other types of words. This is difficult to explain, especially since many children misspelled their own names. It is possible however that personal names in this culture are particularly difficult to spell for beginning spellers, as they tend to be loan words from Arabic or a European language and contain unfamiliar phonemes or phoneme combinations, which in some cases may also be particularly challenging items for speakers of this dialect. In addition, some personal names may have no generally accepted spelling (as in English and other languages), since speakers fluent in another language may spell the name as if it was a word in that language.

For example, the Kiswahili name Hadija is Arabic derived, and a speaker of Arabic – such as a parent or teacher who has some Islamic education – may spell this name KHADIJA to reflect the initial fricative which is present in the Arabic pronunciation but not in the Kiswahili pronunciation.

Place names, however, do not have these ambiguities and in addition may be seen on sign posts, school name boards and in other public forms whereas many other words are not so prominent in the environment. This may in fact be these children's first exposure to print.

There are some other types of errors that were found that may also be due to somewhat different phonological representations to literate adults. Children tended to spell words that should have a double vowel (such as *kaa*) with a single vowel (as KA). Although this word has two syllables, with the stress on the first /*α*/, in casual speech it may not always have a fully articulated pair of vowels, and it may be more difficult for children to represent these words accurately. Indeed, some children spelled a one-syllable word which has a single final vowel (*na*) with two vowels (as NAA), suggesting that they are confused about which of these two short words with only one type of vowel is spelled with a single and which with a double vowel. Words with a single consonant and one type of vowel in Kiswahili can either be one-syllable words, with a single vowel, or two syllable words, with a double vowel. The latter are more common as single syllable words are rare in Bantu languages in general. Children who spell one-syllable words with a double vowel are therefore exhibiting both orthographic and linguistic knowledge. Treiman & Cassar (1997) found children learning to read English had different, and arguably more accurate, phonological representations of

words – for example, they counted three sounds in “hi” whereas adults only counted two.

There were also some differences between the two experiments. In Experiment 1, where children were free to choose the words they spelled, children in higher grades chose to spell more words and those words had more grammatical morphemes per word. This may have led to some increased levels of errors being found, particularly in some grammatical morphemes, in children in higher grade levels in this experiment; children in older grades had an increased tendency to add or drop Y and W in experiment 1, but not in experiment 2, where increasing spelling skill reduced the tendency to add W at the grammatically important morpheme boundary.

It is possible however that the free nature of the task in experiment 1 led the children to perform more poorly on their spelling than might be the case when they are asked to direct their attention to their spelling. Although no significant differences in error rates were found between the two experiments, children in experiment 1 had on average more years of education than children in experiment 2, and should be expected to make fewer errors; either they were choosing more difficult words, or were not paying attention to the accuracy of their spelling due to the free nature of the task.

In Experiment 2 this uncertainty was addressed by analysing errors made in a classroom spelling test, where children were specifically asked to pay attention to their spelling. Although there were very few differences between children in different grades in this task, it is possible that the small range of grade levels (only children from second and third grade were tested), as well as the patchy nature of children’s instructional input, led to little variation between grades compared to the overall variance between children. However the analyses comparing good and poor spellers reveal that some specific types of spelling skill are preferentially picked up when children become better spellers.

Some combinations of phonemes found in Kiswahili are either not found in English and other languages previously studied, or have some phonetic differences. It was found in general that children seem to make the same types of errors despite these differences. In particular, when children are asked to spell consonant clusters consisting of a nasal consonant followed by a stop, fricative or affricate, they tend to make the same types of errors as children learning to spell English (Treiman et al., 1995); namely, they omit or misspell the nasal consonant rather than the second consonant of the cluster. Children learning to spell English tend to spell “hand” as HAD; children learning to spell Kiswahili tend to spell *wengi* “many people” as WEGI. This type of error is particularly interesting as one explanation for the errors found in English has relied on the nasalisation of the previous vowel in words such

as “hand”. It is hypothesised that children see the representation of the nasal as being present in the vowel.

Children make fewer errors of this type in French, where they are specifically instructed to spell nasalised vowels as a vowel followed by a nasal consonant (Treiman et al., 1995). In Kiswahili however this cannot be the cause. In some words there is no previous vowel, as in words with an initial cluster consisting of nasal-other consonant. Where the cluster is word-medial, the preceding vowel is not nasalised.

Hence it can be seen that the process of learning to spell, even in a regularly spelled language, involves the acquisition of many different types of linguistic knowledge. Neither a simple phonological route nor the invocation of more than one parallel routes are adequate for explaining the process; neither does a model in which children pass sequentially through a number of discrete stages. The study of such languages, with regular spellings and a syllable structure that is very different to English, may well help to elucidate some of the processes that children go through when learning to spell. The type of error discussed above – errors in nasal-other consonant clusters of types that cannot exist in English – seems to have much potential, as a means of determining which factors affect children’s, errors of this general type (errors where the phonological realisation of a word appears to be missing one element). The problem for beginning spellers may be that their representation of such phonological elements differs from that of accomplished spellers.

Here we have shown many similarities between Kiswahili and previously studied languages in the types of errors made by beginning spellers, and the reasons for those errors. Kiswahili, in common with other languages (Bryant et al., 1999; Treiman et al., 1995) may well have something more to teach us about the types of skills children need when learning to spell their native language, whatever that language happens to be.

Acknowledgements

We would like to express our thanks to Kathleen Hubbard for linguistic advice, as well as to Fausta Ngowi, Rajabu Mandali, Juliet Mdusi and Christina Mwita of the MAKWAMI project in Bagamoyo, Tanzania for their help in preparing the test materials and assisting in collection of the data reported in this paper. The study was carried out under the auspices of Partnership for Child Development (PCD, Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, Oxford University) and Ushirikiano wa Kumwendeleza Mtoto Tanzania (UKUMTA) – The Tanzania Partnership for Child Development. Support for this research was provided by the James S. McDonnell Foundation, grant reference 94-13.

Note

1. Notation conventions in this paper are as follows: English words, and English translations of Kiswahili words, that children may attempt to spell are in double inverted commas: “cat”. Phonetic renderings of words or phonemes are between slash characters: /kat/, /pɑkɑ/. Kiswahili or other non-English words are in italics: *paka*. Correct spellings or attempts at spelling words or phonemes are in capital letters: KAT, PAKA. So, for example, children may attempt to spell the Kiswahili word *paka* (“cat”, pronounced /pɑkɑ/) as PAK.

Appendix 1 – Phoneme-grapheme correspondences in Kiswahili

Grapheme	Phoneme	As in
A	/ɑ/	Between <i>cat</i> and <i>car</i>
B	/b/	<i>box</i> but implosive
CH	/tʃ/	<i>cheese</i>
D	/d/	<i>dog</i> but implosive
DH	/ð/, dialect /z/	<i>this</i>
E	/ɛ/	<i>pen</i>
F	/f/	<i>fan</i>
G	/g/	<i>gate</i> but implosive
GH	/ŋ/	Voiced velar fricative (<i>loch</i> but voiced)
H	/h/	<i>hat</i>
I	/ɪ/	End of <i>happy</i>
J	/dʒ/	<i>Didya</i> but implosive
K	/k/	<i>kick</i>
L	/l/	<i>leg</i> (not dark L)
M	/m/	<i>mat</i>
N	/n/	<i>nap</i>
NG'	/ŋʲ/	<i>sing</i>
NG	/ŋg/	<i>hunger</i>
NY	/ɲ/	<i>canyon</i>
O	/o/	<i>box</i>
P	/p/	<i>pat</i>
R	/r/	Velar flap
S	/s/	<i>sip</i>
SH	/ʃ/	<i>ship</i>
T	/t/	<i>top</i>
TH	/θ/, dialect /s/	<i>thin</i>
U	/u/	<i>flume</i>
V	/v/	<i>very</i>
W	/w/	<i>wet</i>
Y	/j/	<i>yellow</i>
Z	/z/	<i>zebra</i>

Appendix 2 – Less common errors in experiment 2

Type of error	Mean no of errors per child	Total errors in data set	Mean proportion of errors (of words with this letter pattern)
Final vowel dropped	0.11	21	0.0022
Nasal consonant substituted	0.11	21	0.0023
Internal vowel between consonants omitted	0.10	18	
Nasal consonant added	0.08	15	
Internal vowel omitted next to other vowel	0.07	13	0.0063
/w/ omitted	0.06	12	0.0319
Vowel changed to fricative or liquid	0.05	10	0.0004
Vowel added (other)	0.04	8	
Stop consonant deleted (other)	0.04	8	0.0010
/l/ or /r/ omitted (other)	0.04	7	
Syllable duplicated	0.03	6	0.0002
Stop or fricative consonant doubled	0.03	6	0.0005
/θ/ changed to F	0.03	5	0.0266
Nasal changed to vowel	0.03	5	0.0006
Stop consonant deleted between vowels	0.03	5	0.0022
L or R added (other)	0.02	3	
Fricative deleted between vowels	0.02	3	0.0011
Stop added between two consonants	0.01	2	0.0005
Stop added between two vowels	0.01	2	0.0006
/l/ or /r/ omitted between two vowels	0.01	2	0.0013
Nasal changed to liquid or glide	0.01	2	0.0002
Fricative added between two vowels	0.01	2	0.0010
/s/ changed to TH	0.01	1	0.0011
Stop changed to vowel	0.01	1	0.0001
Vowel changed to stop	0.01	1	0.0000
Stop consonant added (other)	0.01	1	
/l/ or /r/ changed to nasal	0.01	1	0.0005
θ changed to T	0.01	1	0.0053
/ð/ changed to D	0.01	1	0.0053

References

- Alcock, K.J. (2001). *Phonological awareness in children in and out of school*. Presented at British Psychological Society Joint Developmental and Educational Sections Conference, Worcester, UK.
- Alcock, K.J., Nokes, K., Ngowi, F., Musabi, C., McGregor, S. & Mbise, A. et al. (2000). The development of reading tests for use in a regularly spelt language. *Applied Psycholinguistics* 21, 525–555.
- Barry, C. (1994). Spelling Routes (or Roots or Rutes). In G.D.A. Brown & N.C. Ellis (eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27–50). Chichester, UK: John Wiley & Sons.
- Barry, C. & de Bastiani, P. (1997). Lexical priming of nonword spelling in the regular orthography of Italian. *Reading and Writing* 9, 499–517.
- Bryant, P. & Kemp, N. (2003). *Do bees buzz? Rule-based and frequency-based knowledge in learning to spell plural-s*. Child Development.
- Bryant, P., Devine, M., Ledward, A. & Nunes, T. (1997). Spelling with apostrophes and understanding possession. *British Journal of Educational Psychology* 67, 91–110.
- Bryant, P., Nunes, T. & Aidinis, A. (1999). Different morphemes, same spelling problems: Cross-linguistic developmental studies. In M. Harris & G. Hatano (eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 112–133). New York: Cambridge University Press.
- Cassar, M. & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of Educational Psychology* 89, 631–644.
- Castles, A., Holmes, V.M. & Wong, M. (1997). Variations in spelling style among lexical and sublexical readers. *Journal of Experimental Child Psychology* 64, 98–118.
- Ceci, S.J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology* 27, 703–722.
- Cossu, G., Gugliotta, M. & Marshall, J.C. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non parallel processes? Theme Issue: Literacy onset in romance language. *Reading and Writing* 7, 9–22.
- Ellis, A.W. (1993). *Reading, writing and dyslexia: A cognitive analysis* (2nd ed.). Hove, East Sussex, UK: Lawrence Erlbaum Associates.
- Frith, U. (1980). *Cognitive processes in spelling*. London: Academic Press.
- Gill, J.T. (1992). The relationship between word recognition and spelling. In S. Templeton & D.R. Bear (eds.), *Development of orthographic knowledge and the foundations of literacy: A Memorial Festschrift for Edmund H Henderson* (pp. 79–104). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Grunwell, P. (1982). *Clinical phonology*. London: Croom Helm.
- Harris, M. & Giannouli, V. (1999). Learning to read and spell in Greek: The importance of letter knowledge and morphological awareness. In M. Harris & G. Hatano (eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 51–70). New York: Cambridge University Press.
- Inter-territorial Language (Swahili) Committee to the East African Dependencies (1956). *A standard English-Swahili [and Swahili-English] dictionary (founded on Madan's English-Swahili [and Swahili-English] dictionary)*. London: Oxford University Press.
- Jastak, J. & Jastak, S. (1965). *The Wide Range Achievement Test*. Wilmington, DE: Guidance Association.
- Jastak, S. & Wilkinson, G. (1984). *The Wide Range Achievement Test – Revised*. Wilmington, DE: Jastak Association.

- Justicia, F., Defior, S., Pelegrina, S. & Martos, F.J. (1999). Sources of error in Spanish writing. *Journal of Research in Reading* 22, 198–202.
- Lennox, C. & Siegel, L.S. (1994). The role of phonological and orthographic processes in learning to spell. In G.D.A. Brown & N.C. Ellis (eds.), *Handbook of spelling: Theory, process and intervention* (pp. 93–109). Chichester: John Wiley & Sons.
- Mann, V.A., Tobin, P. & Wilson, R. (1988). Measuring phonological awareness through the invented spelling of kindergarten children. In K.E. Stanovich (ed.), *Children's reading and the development of phonological awareness* (pp. 121–147). Detroit, MI: Wayne State University Press.
- Masterson, J. (2000). Developmental surface dyslexia. In E. Funnell (ed.), *Case studies in the neuropsychology of reading* (pp. 123–148). Hove, UK: Psychology Press/Taylor and Francis.
- Nunes, T., Bryant, P. & Bindman, M. (1997a). Morphological spelling strategies: Developmental stages and processes. *Developmental Psychology* 33, 637–649.
- Nunes, T., Bryant, P. & Bindman, M. (1997b). Learning to spell regular and irregular verbs. *Reading and Writing* 9, 427–449.
- Nunes, T., Bryant, P. & Bindman, M. (1997c). Spelling and grammar—The necsed move. In C.A. Perfetti, L. Rieben & M. Fayol (eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 151–170). Mahwah, NJ: Lawrence Erlbaum Associates.
- Oeney, B. & Durgunoglu, A.Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics* 18, 1–15.
- Pinheiro, A.M.V. (1995). Reading and spelling development in Brazilian Portuguese. Theme Issue: Literacy onset in romance language. *Reading and Writing* 7, 111–138.
- Read, C. (1986). *Children's creative spelling*. London: Routledge & Kegan Paul.
- Rego, L.L. & Bryant, P.E. (1993). The connection between phonological, syntactic and semantic skills and children's reading and spelling. Special Issue: Prediction of reading and spelling evidence from European longitudinal research. *European Journal of Psychology of Education* 8, 235–246.
- Snowling, M. (1982). The spelling of nasal clusters by normal and dyslexic children. *Spelling Progress Bulletin* 22, 13–18.
- Snowling, M. (1994). Towards a model of spelling acquisition: The development of some component skills. In G.D.A. Brown & N.C. Ellis (eds.), *Handbook of spelling: Theory, process and intervention* (pp. 111–128). Chichester: John Wiley & Sons.
- The Partnership for Child Development (1999). Short stature and the age of enrolment in Primary School: Studies in two African countries. *Social Science and Medicine* 48, 675–682.
- Thorstad, G. (1991). The effect of orthography on the acquisition of literacy skills. *British Journal of Psychology* 82, 527–537.
- Treiman, R. (1985). Phonemic awareness and spelling: Children's judgments do not always agree with adults'. *Journal of Experimental Child Psychology* 39, 182–201.
- Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology* 83, 346–360.
- Treiman, R. (1993). *Beginning to spell: A study of first-grade children*. New York, NY: Oxford University Press.
- Treiman, R. (1997). Spelling in normal children and dyslexics. In B.A. Blachman (ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 191–218). Mahwah, NJ: Lawrence Erlbaum Associates.
- Treiman, R. & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology* 33, 771–780.

Treiman, R., Goswami, U., Tincoff, R. & Leavers, H. (1997). Effects of dialect on American and British children's spelling. *Child Development* 68, 229–245.

Treiman, R., Zukowski, A. & Richmond-Welty, E.D. (1995). What happened to the “n” of sink? Children's spellings of final consonant clusters. *Cognition* 55, 1–38.

Address for correspondence: Katie Alcock, Department of Psychology, City University, Northampton Square, London, EC1V 0HB, UK
E-mail: k.j.alcock@city.ac.uk