Using Spelling Skills in Brazilian Portuguese and English*

JANE CORREA
MORAG MACLEAN
ELISABET MEIRELES
TANIA LOPES
DORRY GLOCKLING

Abstract

This study examines qualitative differences in children's orthographic knowledge and their ability to develop and test hypotheses about an unknown word in two languages, Brazilian Portuguese and English. A cross-linguistic comparison allows examination of the nature of children's spelling development under different orthographic conditions. 81 Brazilian children in grades 2, 3 and 4 (i.e. aged 9, 10 and 11 years) and 95 English children in the equivalent school grades were asked to play the game 'Hangman' with 4 and 5 letter words. They were also asked to justify their selection of letters. There was greater heterogeneity in the Brazilian children's responses than in the English children's responses across word lengths and syllabic patterns. This is interpreted in terms of the Brazilian children's sensitivity to syllabic patterns. Children's justifications indicated that in all three grades and in both deep and shallow orthographies they made use of lexical and non-lexical strategies.

* The authors are indebted to the schools who welcomed us and the Educational Authorities in Rio de Janeiro and Oxforshire. Support for this research was provided by FAPERJ, MCT/CNPq – Ministry of Science and Technology / National Council for Scientific and Technological Development – and the Oxford Brookes University QR Fund.

Journal of Portuguese Linguistics, 6-2 (2007), 61-82 ISSN 1645-4537

This study aims to gain insight into qualitative differences in children's orthographic knowledge and their ability to develop and test hypotheses about an unknown word. Tests of children's spelling abilities have traditionally been the source of evidence of children's implicit understanding of the nature of their orthography. Researchers make inferences about children's understanding of spelling rules using spelling errors as the data (Treiman, 2004) or by using carefully devised tasks which allow inferences to be made about the use of analogies (Nation & Hulme, 1998) morphological understanding (Nunes, Bryant & Bindman, 1997; Landerl & Reitsma, 2005; Sénéchal, Basque & Leclaire, 2006) or knowledge of legal letter combinations (Cassar & Treiman, 1997). While these approaches have yielded valuable insights into the development of children's spelling, they are school-like tasks, and as such provided a much more formal context in which children have their skills evaluated. Besides that, these tasks explicitly required children to use their orthographic knowledge as an end in itself. From the point of view of children's action, the sole purpose of accomplishing these tasks is to have their spelling skills evaluated. Thus these tasks do not usually tell us much about children's spontaneous use of spelling knowledge. For this we have to look at more qualitative work such as that of Read (1986) on children's invented spelling, Downing, Coughlin & Rich (1986) on children's judgements of correct spelling or Caravolas, Kessler, Hulme and Snowling (2005) on the strategies children report using in response to a spelling task. On the other hand, in non-school contexts children's spelling skills are used as tools for achieving another purposeful goal, such as when children write birthday cards or amuse themselves playing Hangman. The data gathered from more qualitative studies with non-school-like tasks may also contribute to our understanding of the nature of spelling development as well as of children's representations of different aspects of their language, especially phonological ones (Read, 1971; 1986; Ehri & Wilce, 1980).

Weekes (1994) and Castles, Holmes & Wong (1997) describe children's spelling styles in terms of the extent to which the children are lexically reliant (i.e. by means of words) or non-lexically reliant. Their definitions were based on the degree to which children made regularisation or lexicalisation errors in their spellings. A regularisation error occurs when a misspelled word could be pronounced to sound like the target word. On the other hand, if the children misspelled a word by writing another word in its place, it was classified as a lexicalisation error. Their data showed that lexically reliant readers tend to be lexically reliant spellers and non-lexically reliant readers tend to be non-lexically reliant spellers. They suggest that both strategies are available to normally developing children for deployment. Further, the fact that the groups of lexical and non-lexical spellers did not differ in chronological age or reading age suggested that these were indeed styles and strategies and not patterns of development

It is now entirely non-contentious to state that reading and spelling in a deep orthography such as English (containing many irregular words) requires both an assembled spelling strategy (based on sounds) and a lexical strategy (Frith, 1980; Goswami & Bryant, 1990). Indeed, the focus of the debate on reading and spelling strategies is not now on the use of lexical or non-lexical strategies, but on the degree to which shallow orthographies (with few irregularly spelled words) may also demand a lexical strategy. Barry (1992) addressed this question, using a priming methodology to contrast non-word spelling in English, Welsh and Italian. Priming occurs when a word is more rapidly processed due to the previous presentation of the same or a similar word. Barry (1992) found that while English unsurprisingly produced significant lexical priming effects (required to cope with the irregularities), a lexical priming effect was also found for Italian and Welsh. Two major implications can be drawn from these results. First, although the orthographic regularity of certain languages allows an efficient use of assembled spelling, this does not imply that the non-lexical strategy is the sole strategy required. The second implication concerns the functional independence of the word-specific and the assembled spelling procedures to the development of the children's orthographic abilities. It seems that lexical and non-lexical strategies are both needed in deep and in shallow orthographies, although their relative importance might vary in different languages.

There is now a long history of research into the impact of the orthography on children's reading (Bryant & Nunes, 1998) but considerably less on children's spelling. The impetus for the research into reading has been the development of models of learning to read which incorporate phonological awareness. In the case of spelling, the transparency of the orthography – namely the regularity of the grapheme-phoneme correspondences – is also a critical feature. If children are helped to spell by learning about grapheme-phoneme correspondences, then languages with more transparent orthography will prove easier to learn to spell than those with more opaque orthography. A cross-linguistic study will allow us to answer questions about the nature of children's spelling development under different orthographic conditions.

This study aims to compare a relatively shallow and a deep orthography (Brazilian Portuguese and English) using the Hangman game, to find out what impact the orthography has on children's understanding of their language. In the game of Hangman a child has to work out an unknown word given only minimal information about it, knowing only the number of letters it has. This task allows children to make explicit their strategies for spelling the unknown word. The game is played by suggesting possible letters (to fill in the template) or words (to match the template). Children playing the game generally propose letters, filling in the spaces until they are able to infer the correct word or confirm their hypotheses. How then do children select letters to propose? It is possible that they do so on the basis of what they know about the rules of spelling, in other words, using a non-lexical strategy. An alternative strategy for children would be to select letters on the grounds that they knew a word of the appropriate length containing that letter. This strategy depends on easy access to the lexicon, easy generation of vocabulary, with few limits to the choice of words except the word length. The Hangman game therefore allows us to investigate children's knowledge of their language as used in the context of a non-school-like task and to look at some of the factors which might influence children's performance.

The impact of schooling on children's performance in the game is likely to be significant. As children spend time in school and as their literacy skills increase, their written and spoken vocabulary increases. Across a span of three school years improvements in performance are likely to be attributable to education effects. In addition some improvements might also be attributable to increased understanding of how to deploy strategies specifically tailored to achieving success in the game.

In addition to education effects it is also possible to use the game to look for the impact of orthographic-specific expectations. The expectations a reader, speller, or in this case a Hangman player, has about the target word are determined by the individual's age and education (as these are not separable), and knowledge of the orthography. Orthographies vary considerably in terms of the salient features of words. In English word length is associated with numbers of syllables, such that longer words are more likely to be of more than one syllable in length. However it is also the case that words varying in length from one to more than seven letters may all be monosyllables, e.g. 'a'[ə], 'be'[bi], 'sea' [siː], 'deep' [diːp], 'flail' [fleɪl], 'glitch'[glɪtʃ], and 'thought' [θɔːt][1]. Other orthographies have tighter constraints on the relationship between letters and syllables, and Brazilian Portuguese is one of these. A word with seven letters could never be a monosyllabic word in Brazilian Portuguese as the language does not contain a syllable which has more than five elements (Silva, 1999; Azevedo, 2005). As a result, children are likely to make predictions about the word to be identified on the basis of its length, by virtue of the relationship between the word length and syllabic patterns.

Teaching methods are intimately connected with the phonological and orthographic structure of the language. Much literacy instruction in Brazilian Portuguese focuses on orthographic analysis of words (Carvalho, 2005; Rizzo, 1989). In some of these methods, children are taught to read by focusing on the correspondence between spoken and written language in which the syllable is the main unit of phonological analysis (Freire, 1980; Vilas-Boas, 1988). As a consequence of both the nature of the language and the related teaching methods, Brazilian Portuguese speakers are likely to make inferences about the syllabic structure of a word from its length, even when word length varies only from four to five letters. Four-letter words generally have a disyllabic

[1] Phonetic transcription according to the Longman Dictionary of Contemporary English (LDCE). Harlow, Essex: Longman, 2003.

pattern. A four-letter word could fit the expectation of being a CVCV pattern. A CV pattern is a simple and common syllable pattern in Brazilian Portuguese (Matzenauer, 2004; Mezzomo & Menezes, 2001). Adding a letter to make the target a five-letter word has the result that a simple syllabic pattern (made up of a pair of CV syllables) is no longer possible. As a result we might predict that there will be no differences in performance on five-letter words in the game, but that for four-letter words there will be differences related to their deviation from the CVCV pattern. This can be contrasted with the case of English where the difference between a four and a five-letter word tells us little about the syllabic patterns of the words and we predict no differences related to syllabic patterns.

1. Study 1: Brazilian children's orthographic knowledge in the context of Hangman

1.1. Method

1.1.1. Participants

Eighty-one monolingual Brazilian Portuguese speaking children from a primary school located in a low socio-economic status neighbourhood in Rio de Janeiro took part in this study: 32 second graders, mean age of 9 years 4 months; 25 third graders, mean age of 10 years 3 months and 24 fourth graders, mean age of 11 years 3 months. The children were taught to read by a method in which the syllable is the main unit of phonological analysis. To start with, the teacher chooses some disyllables or trisyllables vocabulary in which each syllable has a CV pattern. Generally these are regular words usually taken from their storybooks or spoken language. Initially the children are taught to read and write the selected words. Only then the words are segmented into their constituent syllables. The children are invited to discover which words can be formed by the re-combination of the constituent syllables. Syllabic families are generated for each of the syllables worked with and children are encouraged to discover words that can be written combining the different syllabic families. Once all the linguistic knowledge and activities that can be derived from these key words are fully explored, new words are chosen until all the main aspects of the Brazilian Portuguese orthography have been introduced.

1.1.2. Procedure

The children were interviewed individually by the researchers. They were presented with a Hangman template comprising a set of 4 blanks (representing a four-letter word) and asked to justify their selection and positioning of letters. After each letter selection the children were given feedback about the correctness of the choice. If they selected a 'correct' letter but placed it incorrectly, they were told that "Yes, there is a (letter selected) in this word but it is (correct location)”. Children entered the correct letters, crossed off letters they had selected already and drew in the details of the ‘hanged man’ on a white-board. The children were not given any time constraints or a fixed number of trials to complete the game. The interviews were tape-recorded and transcribed.

Seven words selected from a list of regular, concrete words within the children’s written vocabulary were used in the task. Pre-testing established that the words presented were within the children’s written vocabulary. There were 4 four-letter words and 3 five-letter words, all disyllabic. Only one word ‘cola’ [ˈkɔla][2] (glue) followed the CVCV pattern. Two other words ‘saia’ [ˈsaja] (skirt) and ‘noiva’ [ˈnojva] (bride) had diphthongs. The words ‘obra’[ˈɔbra] (construction site) and ‘fruta’ [ˈfruta] (fruit) had respectively two-consonant final and initial clusters. Finally the words ‘unha’ [ˈuɲa] (nail) and ‘rolha’ [ˈxoʎa] (cork) had digraphs in the last syllable.

1.2. Results and discussion

The mean number of trials used to finish the game was taken as a measure of the level of difficulty children found in solving the task. We computed all the children’s attempts until they guessed the target-word or until all the blank spaces were filled with a letter. Table 1 shows the mean number of trials used by children of different levels of schooling to finish the game. Although the mean number of trials was higher for the five-letter words than it was for the four-letter words, there was less variability between the five-letter words. On the other hand, the mean number of trials to success varied a great deal between the four-letter words, indicating a different level of difficulty for these words.

Table 1. Mean number of trials used by Brazilian children to success in the game

	4 letters				5 letters		
	cola [ˈkɔla]	obra [ˈɔbra]	saia [ˈsaja]	unha [ˈuɲa]	fruta [ˈfruta]	noiva [ˈnojva]	rolha [ˈxoʎa]
2nd grade	9.34	10.47	12.84	14.06	13.22	14.66	14.19
	(4.33)	(5.56)	(4.28)	(5.32)	(3.58)	(4.11)	(3.92)
3rd grade	8.92	9.16	11.96	13.84	11.00	12.24	13.32
	(3.29)	(4.05)	(4.41)	(4.22)	(3.38)	(3.70)	(4.77)
4th grade	6.88	7.63	9.13	13.75	12.08	11.96	12.04
	(3.00)	(3.19)	(2.69)	(4.02)	(3.98)	(4.29)	(5.17)

Note: Standard deviations in brackets.

[2] Phonetic transcription based on Brazilian Portuguese spoken in the city of Rio de Janeiro and surroundings according to the Collins English-Portuguese/Português--Inglês Dictionary, Glasgow: Harper Collins, 2005.

Two separate Analysis of Variance were run for the four- and five-letter words respectively in order to examine differences in children's scores due to schooling or to the syllabic pattern of the word presented or both. The results of these analyses show that there was no significant difference in children's scores for the types of five-letter words presented ($F (2,156) = 1.48$, $p = 0.23$) but that children's performance was affected by the type of four-letter words chosen as a target ($F (3,234) = 27.23$, $p < 0.01$). Post-hoc tests (Newman-Keuls) showed significant differences between 'unha' [ˈuɲa] and all the other words ($p < 0.05$) as well as between 'saia' [ˈsaja] and all the other words ($p < 0.05$). No significant difference was found between 'cola' [ˈkɔla] and 'obra' [ˈɔbra]. The words 'cola' [ˈkɔla] and 'obra' [ˈɔbra] were considered the easiest, followed by 'saia' [ˈsaja]. 'Unha' [ˈuɲa] was considered the most difficult word.

These results are in accord with our predictions and can be explained in terms of the extent to which the syllabic orthographic pattern of the words differed from a CVCV pattern. The fact that there is no significant variation for the five-letter word means is not surprising, as children are frustrated in their expectation of finding a simple CVCV word as the answer. This frustrated expectation complicates matters for them regardless of the actual syllabic pattern of the words presented to them.

When the word has 4 letters the CVCV pattern is a salient and possible choice for the children. The difficulty progressively arises when the syllabic pattern of the word is very different from the CVCV pattern. The word 'cola' [ˈkɔla] obeys the CVCV pattern. Although the word 'obra' [ˈɔbra] does not follow the CVCV pattern, it has high frequency letters and a regular grapheme-phoneme correspondence. The word 'saia' [ˈsaja] caused more difficulty for the children because of the presence of a semivowel instead of the consonant expected if the word was obeying a CVCV pattern. Once the children discovered the second and fourth letter of the target word, they found their expectation of a CVCV pattern confirmed, which made it difficult for them to look for alternative syllabic patterns. Besides not following a CVCV pattern, 'unha' [ˈuɲa] is not a regular word as it contains a digraph (nh). Digraphs lead to mismatches between the numbers of elements in the phonological and the orthographic syllabic patterns.

The results of the tests also show the effect of schooling on children's performance in four-letter word ($F (2,78)=9.23$, $p < 0.01$) and five-letter word ($F (2,78)= 6.80$, $p < 0.01$) trials. For four-letter words, post-hoc comparisons revealed significant differences between the second and fourth graders ($p < 0.05$), and again between the third and fourth graders ($p < 0.05$) in the mean number of trials taken to finish the game. Post-hoc tests (Newman-Keuls) showed that for the five-letter words differences existed between second and third grades ($p < 0.05$) but not between third and fourth grades. These results are interesting as they can lead us to some hypotheses about the strategies that children of different level of instruction are using to solve the task. The decrease in the mean number of trials between grades 2 (mean age 9 years 4 months) and 3 (mean age 10 years 3 months) as compared with grade 4 (mean age 11 years 3 months) for the four-letter words (regardless of word pattern) indicates a significant improvement in children's spelling ability at the end of their primary school years. This improvement might be due to their reading ability, or to increases in their rule based knowledge of the orthography of the language, or to both of these factors.

The fact that a different pattern was observed for the effect of schooling on children's performance in the five-letter trials implies that the strategies children use for the 4-letter words cannot be applied in the same way to five-letter words. Given the difficulty imposed by the five-letter word trials, the major shift in children's orthographic abilities to solve those tasks should occur earlier than it does for the four-letter words. Examination of the children's performance supports this possibility. Inspection of the children's letter choices shows children trying to solve the four-letter word tasks by relying on a lexical strategy, in other words by drawing on their reading and spelling vocabulary and proposing possible words to fit the Hangman pattern. An increase in the number of letters in the word (e.g. from 4 to 5 letters) or recognition (in the case of four-letter words) that the word does not follow the CVCV pattern makes the use of the lexical strategy difficult. A switch to a non-lexically based strategy is then required.

Categorising the children's letter choices as indicating a lexical or non--lexical strategy permits an investigation of the hypothesis that the difference between performance on five and four-letter words across the three school years might be explicable also in terms of differential patterns of strategy use.

1.2.1. Coding strategy use

Children's choices of letters to insert in the Hangman template were coded as based on a lexical strategy if they referred to a possible solution word, which was phonologically appropriate or visually similar to the target word (e.g. "'b', it might be 'bola' [ˈbɔla] "). All other letter choices were coded as non--lexical as they were not word-based choices (e.g. "it might have an 'a' in it").

1.2.2. Opening, development and end-game

As children play the game, they begin with minimal information (only the number of letters in the target word). After a few letter choices they might have both positive and negative information. This information might include knowing about letters that are and are not in the target word, and for letters that are in the target word, the position of those letters with the consequent constraints on future letter choices. By the 'end-game', children will be in a position to propose some likely solutions to the puzzle. Consequently, different parts of the game might lend themselves naturally to differential strategy use. The first two choices can be seen as the 'opening', with the final two choices as the 'end-game'. The 'development' will vary in length depending upon the performance of each child, and will comprise the section of the game between opening and end-game. For each grade, the percentages of all choices that could be classified as lexical were calculated over the number of trials for each phase of the game separately.

Table 2 shows the percentage of letter choices for four and five-letter words which were based on a lexical strategy for each of the three school grades in each of the three phases of the Hangman game. Children in grade 2 (mean age 9 years 4 months) used the lexical strategy more frequently for the four than for the five-letter target words. The lexical strategy is also the children's primary strategy for the opening when the target is a word of 4 letters. However in the development (the mid section of the game), there is a relative decrease in the use of a lexical strategy, possibly due to the constraints imposed by the children's previous letter choices and the limitations of their vocabulary. Spontaneous production of plausible words to fit a partially completed Hangman template is difficult and provoked a shift in the strategy used, i.e. towards an increase in non-lexical strategy. It is clear from Table 2 that it is at the development stage that a non-lexical strategy was dominant for the four-letter word tasks. In contrast, the use of lexical strategy never became dominant over the use of a non-lexical strategy for five-letter words. It seems that the more the syllabic pattern deviates from a CVCV pattern, the more difficult it is for the children to retrieve a word from their vocabulary to help them to solve the task.

Table 2. Percentage of lexical and non-lexical strategies used by Brazilian children

	Phases of the game					
	Opening		Development		End game	
	4 letters	5 letters	4 letters	5 letters	4 letters	5 letters
2nd grade						
Lexical	57	46	33	23	50	43
Non-lexical	43	54	67	77	50	57
3rd grade						
Lexical	51	42	51	27	48	52
Non-lexical	49	58	49	73	52	48
4th grade						
Lexical	59	47	50	36	54	50
Non-lexical	41	53	50	64	46	50

For grade 3 (mean age 10 years 3 months) there is a slightly different pattern in the use of the lexical strategy, in that unlike the second graders (mean age 9 years 4 months), lexical strategy use does not reduce during the development stage of the game for four-letter words. This might be due to a relative improvement in children’s vocabulary by grade 3 (mean age 10 years 3 months). Children can still retrieve a list of possible words from the lexicon despite the constraints arising from their previous choices. The fact that they can do this is also shown by the increase in the frequency of lexical strategies used in the ‘end-game’ for five-letter words as compared to second graders. Although an increase in the children’s vocabulary does not initially help them to solve the game quickly when the word has more than 4 letters, it proves to be highly useful for the end-game. The improvement in children’s vocabulary allows them to take advantage of their previous letter choices to narrow the range of plausible words to complete the game.

Comparing fourth graders with third graders resulted in relatively small differences, the most noticeable effect being mainly related to the increase of their lexical strategy use in the game opening. This is again largely explicable in terms of an improvement in children’s vocabulary over the school years. The lexical strategy must be based on children’s reading and spelling and spoken vocabulary. Although this strategy is important for the solution of the game for the children of all grades, it is not always possible for the children to deploy it, especially when the target word increasingly deviates from the CVCV pattern. This poses the question of what the use of a non-lexical strategy tells us about the children’s reasoning.

The analysis so far indicates that the children are approaching the target word by means of letter-by-letter assembly of the word. It does not tell us the extent to which the children’s choices are based on their understanding of the orthographic rules of their language. Are their non-lexical choices indicative only of simple assembly of the target word using letters of high frequency, presented in highly frequent sequences such as vowel strings or using alphabetical order? A qualitative examination of the protocols is required to answer this question.

The examination of the children’s use of the non-lexical strategy showed us that their choices can be made on the basis of orthographic rules but this might not always be the case. Children’s knowledge about the orthographic pattern of the words is mainly related to the requirement for vowels and the constraints of some letter sequences.

In relation to vowels, as can be seen from protocol A, children acknowledge their importance as the centre of the syllable, generally beginning the game with a vowel because “all words have a vowel”. The vowels themselves were not chosen at random, but according to their frequency. Children started with ‘a’, followed by ‘o’ or ‘e’, then ‘i’, with ‘u’ as their last choice. Generally children stop choosing vowels when they get two of them right.

1.2.3. Protocol A: Focusing on vowels

_ _ _ A	A	Almost all the words have the letter A[3]
C a _ a	C	I think it is ‘casa’ [ˈkaza] (house)
P a _ a	P	It can be ‘para’ [ˈpara] (for)
_ a M a	M	Because it can be ‘cama’ [ˈkama] (bed). Oops, I forgot the word does not have a C
S a _ a	S	It is ‘sala’ [ˈsala] (living room)
s a L a	L	It is ‘sala’! [ˈsala]
s a I a	I	I was thinking of the vowels which I did not put because I could not find any consonant which fitted there, making a word that I knew

Once the children have some clues, combinations of letters have been suggested either because of specific positions in the word or because of the presence of other letters. Protocol B illustrates this in the case of ‘unha’ [ˈuɲa].

1.2.4. Protocol B: Using letter combinations/positions

u N h a	N	Because I thought of those letters that can go with H. I thought of an L, but it wasn’t it. It was not the C either. Then it could only be the N. That is: unha [ˈuɲa].

The importance of considering position and frequency of the letters for an appropriate choice was explicitly stated by some children: “Every time I play this game, I start with the vowels. There are always vowels in words. It is easier. Almost in all the words the second (letter) is a vowel.”

Children cannot always explain their reasoning or show a clear rationale for their choices. However this does not mean that those choices are random. Some of their choices show their sensitivity to orthographic features, in particular to letter frequency. Only under certain circumstances are children’s choices apparently random. Those are the cases when they have already had many attempts and still do not have any idea about the identity of the word. They generally get exasperated about the possibility of losing the game, and then try all sorts of letters, as shown by Protocol C.

1.2.5. Protocols C: Using alphabetical order

_ o I _ a	I	Because an I is a very well known letter, like an A and an E
D o i D a	D	I thought of doida [ˈdojda] (crazy)
_ o i G a	G	I am going in alphabetical order
_ o i J a	J	I am going in alphabetical order
_ o i L a	L	I am going in alphabetical order

3 Upper case letters in the protocols indicate the letter choice for the turn, lower case letters indicate letters identified in previous turns.

M o i _ a	M	I've picked the end so many times that now I'm going to pick the beginning. I am going in alphabetical order
N o i _ a	N	Ah, now I can see the word, it is noiva [ˈnojva].

It is worthwhile to point out that the use of the alphabetical order does not necessarily mean a irrational choice. Children can use it as an organised way of getting to the correct word: "It is cola [ˈkɔla]. (How do you know that?) Because of the letters. I went through the alphabet, skipping some letters because they could not make it sound right."

In the case of five-letter words and the difficulty they impose by virtue of not fitting the CVCV pattern, some children felt a need to mention this in their explanation for their choices. They showed sensitivity to the fact that they would come across consonant clusters, diphthongs or digraphs: "Generally when there are five (letters) there might be loads of consonants. Then I keep putting vowels at the end because it makes it easy… (pause)… well, sometimes there are lots of vowels in the middle as well."

1.3. Conclusions

Observing children solving a task such as Hangman gives us a valuable insight into their reasoning about spelling in Brazilian Portuguese. Study 1 has shown that two main strategies play an important role in the children's success in solving the task. The first one, a lexical strategy, allows children to retrieve a range of plausible solution words based on available information. This strategy depends mainly on children's vocabulary. As a result, its effectiveness relies on the improvement of the child's vocabulary and increases in the speed of access to that vocabulary. A non-lexical strategy is used selectively by children when the word appears unlikely to follow the CVCV pattern. However, the use of a non-lexical strategy does not always mean that children are aware of the basis of their letter choices. For example, a child might not produce any justification of their choice but still select high frequency letters indicating some knowledge of the constraints imposed within the game. Explicitly rule-based choices would be indicated by children showing awareness of the constraints of the partial solution or word length, or by their sensitivity to letter-string frequencies or other orthography-specific knowledge. A different orthography, a deeper one such as English, might produce a different pattern of results, specifically in the response of children to four and five-letter words, and possibly also in the rationales for their letter choices.

2. Study 2: English children's orthographic knowledge in the context of Hangman

2.1. Method

2.1.1. Participants

Ninety-five children from primary schools in Oxfordshire, England took part in this study: 32 in year 4, mean age of 8 years 11 months; 31 in year 5, mean age of 9 years 11 months and 32 in year 6, mean age of 10 years 11 months. They were all monolingual, being English their native language. The children were taught to read and spell by the whole-language approach.

2.1.2. Procedure

The procedure followed was similar to that of Study 1. Children were interviewed individually by the researchers. They were presented with either three or four words drawn from a list of 17 concrete words. Pre-testing established that the words presented were within the children's written vocabulary. All words were of either four or five letters. All four-letter words were all monosyllabic, following CVCC, CVCV, CCVC or CVCC patterns (e.g. sock [sɒk], cake [keɪk]). The five-letter words were either monosyllables (CVCCV, CCVVC, CCVCC, e.g. drink [drɪŋk], steam [stiːm]) or disyllables (CCVCC, CVCVC e.g. salad [ˈsæləd], gravy [ˈgreɪvi]). All children were presented with at least one four-letter word and at least one five-letter word. The consonant vowel patterns in the words were selected as highly frequent in four and five--letter words that are high on concreteness and imagery ratings according to the MRC Psycholinguistic Database (Coltheart, 1981). The Hangman game was played in an identical fashion to the game in Study 1. Children's choices and justifications were tape-recorded and transcribed.

2.2. Results and Discussion

As in Study 1, the number of trials to finish each game was calculated. Table 3 shows the mean number of trials taken by children of different levels of schooling to complete the game for monosyllabic four and five-letter words as well as disyllabic five-letter words. While individual words were presented as the target in the game to small numbers of children, making an analysis of variance by word impracticable, it was clear from the mean number of trials taken to identify the words that there was little variation in the difficulty of identifying the words.

Table 3. Mean number of trial used by English children to success in the game

	4 letters	5 letters	
	One syllable	One syllable	Two syllable
Year 4	12.14	11.71	12.82
	(3.37)	(3.14)	(5.53)
Year 5	10.87	12.08	11.28
	(3.37)	(3.16)	(4.21)
Year 6	11.00	10.38	11.28
	(3.46)	(4.08)	(3.51)

Note: Standard deviations in brackets.

Three separate Analyses of Variance were run for the four-letter words and the two types of five-letter words (monosyllabic and disyllabic) in order to examine differences in children's scores due to schooling or to the syllabic pattern of the word presented or both. No significant differences were found between the grades for four-letter words (F (2,92) = 1.35, p = 0.264), five-letter monosyllabic words (F (2,76) = 1.62, p = 0.205) or five-letter disyllabic words (F (2,61) = 0.80, p = 0.453). The results show neither significant differences between the word types nor any schooling effect. The lack of effect of schooling is puzzling as one might expect children's performance to improve with age or spelling experience. However, in English, given the range of CV patterns permissible in four and five-letter words, it is the case that even identifying one letter, for example finding a vowel early in the game, does not limit choice to any very great degree. Even the identification of a vowel and a letter for a short English word still leaves many possibilities. As a result, it is not surprising that the average number of trials to completion is only just below half of the alphabet length.

The use of strategies to complete the game may provide further clues as to the lack of schooling effect. The children's choices were coded using the same lexical/non-lexical categories as used in Study 1. Table 4 shows the percentage of lexical and non-lexical choices made in each year group for each phase of the game. The percentages were calculated as for in the Study 1. For each school year the pattern of strategy use was largely similar; children were more likely to use lexical strategies for four than for five-letter words. However, very few lexical choices are made in the opening and development phases of the game (never making up more than 26% of the choices), while the end-game is dominated by lexical choices with over 58% of all choices being lexical for all three school years. Where there are differences between the school years they lie in the marginally greater readiness of Year 4 (mean age 8 years 11 months) children to make lexical choices in the opening when compared with children in Years 5 (mean age 9 years 11 months) and 6 (mean age 10 years 11 months).

These results are of interest in that they do not follow the pattern usually expected in spelling tasks in English where a considerable proportion of lexical strategies have been found. One reason for this may lie in the nature of the game. The Hangman game for English speakers gives little distinguishing information to children in the early stages. A four or five-letter word might be of one or two syllables, and might conform to several high frequency CV patterns. Until several letters are in place the demands on the child's vocabulary in using a lexical strategy may simply be too heavy for this to be feasible. It is easier to draw on the alphabet – which after all has only 26 letters as opposed to the 20,000 thousand strong traditional vocabulary of Shakespeare's English. If the game-player is not in danger of being 'hanged' too soon, then a non-lexical, assembly-by-letter process may indeed be the best strategy. However, by the end-game the range of potential words has narrowed, and children were able to use the lexicon to propose a series of possibilities until they had correctly identified the word.

Table 4. Percentage of lexical and non-lexical strategies used by English children

	Phases of the game					
	Opening		Development		End game	
	4 letters	5 letters	4 letters	5 letters	4 letters	5 letters
Year 4						
Lexical	26	12	19	9	70	61
Non-lexical	74	88	81	91	30	39
Year 5						
Lexical	13	9	23	12	61	59
Non-lexical	87	91	77	88	39	41
Year 6						
lexical	15	10	18	9	65	59
Non-lexical	85	90	82	91	35	41

Qualitative analysis of the protocols of the English children illustrates their rationales. Lexical strategies tended to be rare, but a few children used this strategy almost to the exclusion of all other strategies. Protocol D illustrates the possibility of using a lexical strategy in a reasoned way. The child's early choices of letters include a vowel and his later suggestions are variations of a theme, using rhymes (MITE [maɪt] and KITE [kaɪt]) and half-rhymes (BATE [beɪt] and CAKE [keɪk]).

2.2.1. Protocol D: Lexical Strategy

N _ _ _	N	Might be nope [nəʊp]
T _ _ _	T	Might be that [ðæt]
D _ _ _	C	Might be date [deɪt]
M _ _ _	M	Might be mite [maɪt]
E _ _ _	E	Might be eel [iːl]
F _ _ e	F	Might be fate [feɪt]
S _ _ e	S	Might be shoe [ʃuː]
B _ _ e	B	Might be bate [beɪt]
K _ _ e	K	Might be kite [kaɪt]
C _ _ e	C	Might be cake [keɪk]

Non-lexical strategies vary from the apparently random to the highly organised. The apparently random approach is exemplified in the following example from year 4 who selects letters apparently almost by lexical means –

thinking of a word that begins with a specific letter. It is far from clear whether the letter or the word came first, but the researcher had the strong impression that in this case the letter came first and the justification was a creative and at times imaginative summoning of salient personal information triggered by the letter. In this example it is interesting to see that the justification for the 'N' appears to be alliterative, including 'next-door' [nekst-dɔː(r)], 'neighbour' [ˈneɪbə(r)] and 'nuts' [nʌts].

2.2.2. Protocol E: Letters with personal justifications for each choice

_ a J _	J	My name starts with J
_ a S _	S	Dad's name starts with S
C a _ _	C	My favourite food is chocolate
_ a N _	N	Next door neighbour is allergic to nuts
_ a n E	E	Mum's favourite food is egg on toast. Most words end with 'e' like time

The non-lexical, assembled-letter strategy is the most frequent one used by all age groups in this study. Children call on their knowledge of successful strategy – identifying vowels first, then moving to high frequency letters, and often calling on their knowledge of plausible positions for letter combinations. This can range from knowing that an 'S' denotes a plural at the end of words, to knowing about specific word endings such as 'ED' (incidentally less likely in short words than longer ones), or the vowel combination of 'EA'. One justification addressed vowel combinations explicitly when the child chose and 'E' after choosing an 'A', saying "They go together to make an 'ee' sound". Larger units can also be identified and these include identifying the silent 'e' when an 'R' is selected after and 'A' blank 'E' have been selected. The child went on to justify the choice of an 'R' by saying "To make an –'are' sound at the end of the word".

2.2.3. Protocols F: Letter combinations and reference to letter position

D _ _ _ _	D	There are lots of words like 'dea' in the beginning
_ E _ _ _	E	It's a popular letter. Lots of words have 'ea'.
_ A _ _ _	A	Lots of words have 'a'
_ _ T _ _	T	I guessed
_ _ _ L _	L	Lots of words have L and E as their spelling
_ _ _ _ S	S	Often at the end for more than one (i.e., plural)

2.3. Conclusions

Observations of English readers and spellers playing the game of Hangman suggest that the strategies they use to play the game result in fairly even performance across years 4 (mean age 8 years 11 months), 5 (9 years 11 months) and 6 (10 years 11 months). Children rely on one main strategy, a non-lexical strategy of assembly by letter until the word possibilities are so limited that a lexical strategy is feasible. In the case of English this does not generally occur in these age groups until the end of the game when almost all letters have already been identified. The letter-by-letter approach may be more or less informed by their knowledge of spelling rules and is likely to be the strategy of choice where the word length gives no indication of the constraints on the word's spelling or syllabic construction.

3. General Discussion

The two studies described above are an example of a task with apparently identical cognitive demands being presented to children developing their literacy skills in two contrasting linguistic contexts. The contrast here is between a deep and a relatively shallow orthography. An identical game results in quantitative and qualitative findings showing interesting differences and similarities in children's reasoning about their orthography.

Children in both linguistic communities completed the tasks (averaged across words) in a similar number of trials. However, for the English children there was little variation between the word types or individual words, whereas for the Brazilian children the four-letter words were completed in fewer trials than the five-letter words, with the notable exception of the case of a word deviating markedly from the CVCV pattern. In other words, there was greater heterogeneity in the Brazilian children's responses than in the English children's responses.

The comparisons of children's ease/difficulty in completing the tasks between the three school years yielded interesting results. For the Brazilian children, schooling effects can be seen in the improved performance of the older children, although this effect is slightly different for the two types of material (i.e. word lengths). In contrast, no such schooling effect could be found for the English children.

The explanation for differential schooling effects must lie with two factors: the constraints of the orthography and the demands of the Hangman game. The Hangman game requires the children to do more than simply spell a word, they have to either generate clues to the identity of the word by proposing letters, or they have to test specific hypotheses which might relate to either a word or a combination of letters. The opening of the game gives them merely the number of letters in the unknown word. This information can provide Brazilians with both the number of syllables in the word and the approximate nature of the intrasyllabic pattern (i.e. a four-letter word can be two syllables each of CV pattern while a five-letter word clearly cannot follow this pattern). It is not surprising that Brazilian children are progressively able to take advantage of this information with increased schooling. No such advantage is conferred on the English children for whom the word length carries no useful information about the CV pattern of a word. Subsequent letter choices are also of limited usefulness in hypothesis testing as the presence of many irregularly spelled words in English means that there are relatively few constraints on their choices. Further, for English children there are no clues to how to segment the word, which could be split into letter clusters at any point (Treiman, 1992).

These findings also highlight the influence of the rhythm of spoken language on children's conception of writing. In stress-timed languages, syllables vary considerably in length and complexity, whereas syllable-timed languages tend to have relatively simpler syllable structure displaying fewer syllable patterns, few clusters and a dominance of the CV syllabic pattern (Ramus, Nespor & Mehler, 1999; Vigário, Frota & Freitas, 2003). It is not always easy to classify a language into one of these two categories as most languages show aspects of both types of timing (Roach, 1982). Languages, however, can be distinguished in terms of the type of timing predominates (Roach, 1982). English could be classed as predominantly stress-timed (Ramus, Nespor & Mehler, 1999), whereas Brazilian Portuguese exhibits mixed rhythmic properties with a relatively high-degree of syllable-timing (Barbosa, 2000; Bisol, 2000).

Radebaugh (1985) found that children's reports of how they spelled a word included strategies such as segmenting words. The nature of the Hangman task makes it very difficult to know where to start the segmentation process until the game is nearly completed. Besides the high inconsistency of phoneme-grapheme correspondences, English exhibits a complex syllabic structure with many consonant clusters and closed syllables (Macizo & Van Petten, 2007). These features make syllable recognition difficult in print. It seems that syllables might not function as sublexical cues to help English children in visually perceiving or retrieving a word from their mental lexicon (Macizo & Van Petten, 2007). English children are therefore disadvantaged in deploying the knowledge that they might possess about spelling rules. For Brazilian children, there is a starting point to the segmentation process based on the basic properties of syllabic structure trigged by the rhythm of Brazilian Portuguese. A four-letter word may not be a CVCV word, but this may be a reasonable initial hypothesis that can be tested. The syllabic pattern of the word is therefore open to investigation. For five-letter words they know immediately that the word is not a CVCV pattern and they expect to find a consonant cluster or a diphthong. Again, these possibilities can be explored with greater efficiency with increased schooling.

Turning to the qualitative analysis of the children's protocols, it is clear that in both groups children's choices of letters for the Hangman game seem to be guided by their sensitivity to the phonology and orthography of their language. Letter choice followed by a justification can be said to show explicit awareness of some phonological and orthographic constraints. However choices without proffered rationales may also show some understanding of phonological and orthographic constraints. In following the children's choices throughout the course of a game, it is clear that both Brazilian and English children are sensitive to certain phonological and orthographic features. First among these is the requirement for words to contain vowels. The relative frequency of consonants is also a feature used by the children in their hypothesis testing.

Many children propose high frequency letters as soon as they have established the identity of the vowels in the words. In addition to this, the protocols show that children are often also sensitive to the frequency of certain letter combinations, although these combinations are language specific.

Lexical and non-lexical strategies were deployed by both Brazilian and English children, both groups making more use of lexical strategies for four than five-letter words. However the pattern of use was markedly different. English children made little use of lexical strategies except in the end-game, and this did not change with increased schooling. Given the inconsistency of the letter-sound mapping and the high variability of syllable patterns in English, children needed the information provided by great number of letters before deeming it worth adopting a lexical strategy in Hangman game. In contrast the Brazilian children showed a much greater use of lexical strategies in the opening and development phases of the game, and showed a pattern which changed with schooling. As Brazilian Portuguese shows a relatively high-degree of syllable-timing (Barbosa, 2000; Bisol, 2000), syllables might function as sublexical units to Brazilian children. Cued by the rhythm of Brazilian Portuguese and focusing on the presence of the vowels, Brazilian children assumed the dominance of CV syllabic pattern for 4-letter word. This assumption guided lexical decision and word production, favouring the use of a lexical strategy in different phases of the game. Our findings support Weekes (1994) and Castles, Holmes and Wong (1997)'s conclusions that both strategies are available to normally developing children for deployment and that they cannot be considered a pattern of development. However, our data do not corroborate their suggestion that those strategies can be a style. Lexical and non-lexical strategies were used in the game by children according to some constraints such as the number of letters in a word.

The use of a lexical or non-lexical strategy is not only related to the nature of the language but also to the cognitive demands of the task as they are shown by the different phases of the game. This finding is an interesting one as it suggests that adherence to the notion that English orthography with its irregularities provokes a lexical response to spelling may be appropriate in the case of spelling tasks, but not for all tasks using spelling knowledge. The counterpoint of this is that a lexical strategy was deployed in a relatively shallow orthography such as Brazilian Portuguese.

Our two studies provide support from qualitative work for Barry's (1992) suggestion that lexical and non-lexical strategies were not independent pathways to spelling. Lexical and non-lexical strategies are used in both shallow and deep orthographies. Our results certainly show that both strategies are used by children of varied ages and stages of their schooling, and in various points of the Hangman game.

4. Implications for Practice

Spelling refers to more than forming a word from letters in a proper way. Spelling can be considered a creative linguistic process which involves children's knowledge of the conventions used in the writing system of their language such as acceptable letter sequences, syllable structure and the ways a phoneme may be represented according to its position in a word (Downing, Coughlin & Rich, 1986; Treiman & Cassar, 1997; Treiman, 1998). Tests of children's spelling abilities give us evidence of children's implicit understanding of the nature of their orthography. However the examination of children's spelling skills in a non-school-like task such as Hangman can give us some evidence of the way children use their orthographic knowledge out of school.

The Hangman task may also be an interesting instrument to analyze both children's strategies and the knowledge they are able to draw on to help them to spell. We believe that empirical evidence from the study of children's spelling strategies using Hangman task which takes into account children's vocabulary level or differences into measures of morphosyntactic awareness can make an important contribution to a theoretical model whose aim is to incorporate a more contextually sensitive theory and method into the analysis of the process by which children learn how to read and write.

We also believe that Hangman task can provide data in a form that may be useful when assessing the impact of teaching on children's reasoning about orthography. The task may offer useful information to teachers assessing children's understanding of the phonological and orthographic patterns of their language, giving access to the process of spelling and thinking about spelling rules. Finally the Hangman task can provide us with valuable insight into children's explicit knowledge which could indicate useful intervention measures at a crucial stage in the development if literacy skills.

References

Azevedo, M. M. (2005) *Portuguese: a linguistic introduction*. Cambridge: Cambridge University Press.

Barbosa, P. A. (2000) Tempo-silábico em Português do Brasil: uma crítica a Roy Major. *D.E.L.T.A.*, 16: 369-402.

Barry, C. (1992) Interactions between lexical and assembled spelling. In *Psychology, spelling and education* (C. Sterling & C. Robson, Editors), pp. 71-86. Clevedon, England: Multilingual Matters.

Bisol, L. (2000) O troque silábico no sistema fonológico (um adendo ao artigo de Plínio Barbosa). *D.E.L.T.A.*, 16: 403-413.

Bryant, P. E. & Nunes, T. (1998) Learning about the orthography: A cross-linguistic approach. In *Global prospects for education. Development, culture, and schooling* (S. G. Paris & H. M. Wellman, Editors), pp. 171-192. Washington, DC: American Psychological Association.

Caravolas, M., Kessler, B., Hulme, C. & Snowling, M. (2005) Effects of orthographic consistency, frequency, and letter knowledge on children's vowel spelling development. *Journal of Experimental Child Psychology*, 92: 307-321.

Carvalho, M. (2005) *Alfabetizar e letrar: um diálogo entre a teoria e a prática*. Petrópolis: Vozes.

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.

Coltheart, M. (1981) The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A: 497-508.

Downing, J., Coughlin, R. M. & Rich, G. (1986) Children's invented spelling in the classroom. *The elementary School Journal*, 86 (3): 295-303.

Ehri, L. & Wilce, L. (1980) The influence of orthography on reader's conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, 1: 371-385.

Freire, P. (1980) *Educação como prática da liberdade*. Rio de Janeiro: Paz e Terra.

Frith, U. (1980) Unexpected spelling problems. In *Cognitive processes in spelling* (U. Frith, Editor). London: Academic Press.

Goswami, U. & Bryant, P. (1990) *Phonological skills and learning to read*. London: Lawrence Erlbaum Associates.

Landerl, K. & Reitsma, P. (2005) Phonological and morphological consistency in the acquisition of vowel duration spelling in Dutch and German. *Journal of Experimental Child Psychology*, 92: 322-344.

Macizo, P. & Van Petten, C. (2007) Syllable frequency in lexical decision and naming of English words. *Reading and Writing*, 20: 295-331.

Matzenzuer, C. L. B. (2004) Bases para o entendimento da aquisição fonológica. In *Aquisição Fonológica do Português* (R. R. Lamprecht, G. F. G. Bonilha, G. C. M. de Freitas, C. L. B. Matzenauer, C. L. Mezzomo, C.C. Oliveira & L. P. Ribas), pp. 33-58. Porto Alegre: Artmed.

Mezzomo, C. & Menezes, G. (2001) Comparação entre a aquisição da estrutura da sílaba no português brasileiro (PB) e no português europeu (PE). *Letras de Hoje*, 36: 691-698.

Nation K. & Hulme C. (1998) The role of analogy in early spelling development. In *Reading and Spelling: Development and disorders* (C. Hulme & R. M. Joshi, Editors), pp 433-445. Mahwah, NJ, USA: LEA.

Nunes T., Bryant P. & Bindman M. (1997) Spelling and grammar – the necsed move. In *Learning to Spell: Research, theory, and practice across languages* (C.A. Perfetti & L. Rieben, Editors), pp. 151-170. Mahwah, NJ, USA: LEA.

Radebaugh, M.R. (1985) Children's perceptions of their spelling strategies. *The Reading Teacher*, 38 (6): 532-536.

Ramus, F., Nespor, M. & Mehler, J. (1999) Correlates of linguistic rhythm in speech. *Cognition*, 73: 265-292.

Read, C. (1971) Preschool children's knowledge of English phonology. *Harvard Educational Review*, 41: 1-34.

Read, C. (1986) *Children's creative spelling*. London: Routledge & Kegan Paul.

Rizzo, G. (1989) *Os diversos métodos de ensino da leitura e da escrita*. Rio de Janeiro: Francisco Alves.

Roach, P. (1982) On the distinction between 'stress-timed' and syllable-timed' languages. In *Linguistic controversies* (D. Crystal, Editor), pp. 73-79. London: Edward Arnold.

Sénéchal, M., Basque, M.T. & Leclaire, T. (2006) Morphological knowledge as revealed in children's spelling accuracy and reports of strategies. *Journal of Experimental Child Psychology*, 95: 231-254.

Silva, T. C. (1999) *Fonética e fonologia do português*. São Paulo: Contexto.

Treiman, R. (1992) The role of intrasyllabic units in learning to read and to spell. In *Reading acquisition* (P. B. Gough, L. Ehri & R. Treiman, Editors), pp. 65-106. Hillsdale, NJ: Lawrence Erlbaum Associates.

Treiman, R. (1998) Beginning to spell in English. In *Reading and spelling: Development and disorders* (C. Hulme & R. M. Joshi, Editors), pp. 371-393. Mahwah, NJ: Erlbaum.

Treiman, R. (2004) Phonology and spelling. In *Handbook of children's literacy* (P. Bryant & T. Nunes, Editors), pp. 31-42. Dordrecht, the Netherlands: Kluwer.

Treiman, R., & Cassar, M. (1997) Spelling acquisition in English. In *Learning to spell: Research, theory, and practice across languages* (C. A. Perfetti, L. Rieben & M. Fayol, Editors), pp. 61-80. Hillsdale, NJ: Erlbaum.

Vigário, M., Frota, S. & Freitas, M. J. (2003) From signal to grammar: rhythm and the acquisition of syllable structure. *Proceedings of the 27th Annual Boston University Conference on Language Development*. Dommerville, Mass.: Cascadilla Press, 809-821.

Vilas-Boas, H. (1988) *Alfabetização: outras questões, outras histórias*. São Paulo: Brasiliense.

Weekes, B.S. (1994) Spelling skills of lexical readers. *British Journal of Psychology*, 85: 245-257.

Jane Correa, Elisabet Meireles
& Tania Lopes
Universidade Federal do Rio de Janeiro
Av. Pasteur, 250, Urca, RJ, Brasil
CEP 22290-240
janecorrea@rjnet.com.br
betmeireles@ig.com.br

Morag MacLean &
Dorry Glockling
Oxford Brookes University
Headington, Oxford, OX3 0BP
United Kingdom
mmaclean@brookes.ac.uk