Scandinavian Journal of Psychology, 2005, 46, 263–272

Orthography as a handicap? A direct comparison of spelling acquisition in Danish and Icelandic

HOLGER JUUL[1] and BALDUR SIGURDSSON[2]

[1]*Department of Scandinavian Research, University of Copenhagen, Denmark*
[2]*Iceland University of Education, Reykjavík, Iceland*

Juul, H. & Sigurdsson, B. (2005). Orthography as a handicap? A direct comparison of spelling acquisition in Danish and Icelandic. *Scandinavian Journal of Psychology*, *46*, 263–272.

Spelling of cross-linguistically very similar nonwords was compared in 115 Danish and 77 Icelandic children (primarily 3rd and 4th graders). Danish children made more errors than Icelandic children on word medial consonant doublets and on word initial consonant clusters, even when the groups compared were matched on simpler spelling tasks. These results suggest that the acquisition of phonemic encoding skills is slower in "deep" orthographies such as Danish than in more "transparent" orthographies such as Icelandic. The effect of orthography was expected for consonant doublets because of the relatively more complex sound-letter correspondences in Danish. For consonant clusters, however, sound-letter correspondences are perfectly regular in both languages. The study thus points to the conclusion that even the mastery of regular sound-letter correspondences may be delayed in deep orthographies.

Key words: Spelling, orthographic depth, consonant doublets, consonant clusters, Danish, Icelandic.

Holger Juul, Centre for Reading Research, Department of General and Applied Linguistics, University of Copenhagen, Njalsgade 80, DK-2300 Copenhagen S, Denmark. Tel: +45/ 35 32 91 32; fax: +45/ 35 32 85 12; e-mail: juul@hum.ku.dk

INTRODUCTION

In recent years there has been a growing interest in cross-linguistic comparisons of literacy skills. In a Nordic context, the IEA-study of the early 1990s that compared reading comprehension skills in children from more than 30 countries (Elley, 1992) received much attention for at least two reasons. First, it showed that reading development is remarkably swift in children learning to read the highly regular Finnish orthography. Second, it showed that children learning to read the much more irregular Danish orthography were significantly behind their peers in all the other Nordic countries. More recent cross-linguistic comparisons have indicated that even in Danish adolescents and adults reading comprehension skills are poorer than in the other Nordic countries (Allerup, Mejding & Zeuner, 2001; UNICEF, 2002).

The relatively poor performances of the Danish participants in these studies may have several explanations. One possibility is that the irregularities of Danish orthography tend to make decoding less efficient (Juul, 2001). Sound-letter correspondences in Danish are often ambiguous, and silent letters abound. The relations between sounds and letters in the other Nordic orthographies are relatively more transparent.

However, one may also point to other differences between Denmark and the other Nordic countries that could have an effect on literacy skills. Although the Nordic countries are very similar on most socio-economic parameters, differences in teacher education, educational system, parental attitudes to literacy achievement, etc., may put Danish readers at a disadvantage (Sommer, Lau & Mejding, 1996). One may even ask whether technical problems due to orthographic irregularities have any practical importance worth talking about beyond the earliest phases of literacy acquisition (Lundberg, 1997). That is the question we address in this paper.

The effects of orthographic depth

Irregularity or complexity in an orthography, that is, deviations from simple and predictable one-to-one sound-letter correspondences, is often referred to as orthographic depth. Studies of the effects of orthographic depth have primarily been concerned with the reading process. An important issue has been whether the lower accessibility of phonological information in deep orthographies leads to a relatively greater reliance on visual-orthographic information in naming and lexical decision tasks (the orthographic depth hypothesis, cf. Katz & Frost, 1992; Frost, 1992).

A second issue has been whether orthographic depth affects the size of units used in phonological recoding (the grain size hypothesis). Recent cross-linguistic studies suggest that reading a deep orthography like English tends to involve larger units than reading a more "transparent" orthography like German (Ziegler, Perry, Jacobs & Braun, 2001; Goswami, Ziegler, Dalton & Schneider, 2003).

A third issue has been whether orthographic depth impedes the development of accurate and fluent reading skills. The present study primarily addresses this third issue. Previous studies that have dealt with this issue have had a natural focus on decoding skills. This is because orthographic depth is presumed to affect decoding directly, while

 ISSN 0036-5564.

only indirect effects would be expected for reading comprehension. If cross-linguistic differences cannot be found for decoding, there will be little reason to suspect that orthographic depth can help explain the differences in reading comprehension that have been found in the international studies referred to above. However, encoding skills should be just as likely as decoding skills to be affected by orthographic depth. Yet, as will become apparent from the review in the following section, few previous cross-linguistic studies have focused on spelling skills. It is also noteworthy that in the majority of previous studies English has been the deep orthography which was compared to other, more transparent, orthographies.

In the present paper we report a cross-linguistic study which focused on spelling acquisition, and in which the deep orthography was not English, but Danish. Danish was compared to another Nordic orthography, namely the relatively more transparent Icelandic orthography. The focus of the study was the spelling of some very specific phonological structures in pupils from Grades 3 and 4. The purpose of the comparison was to test the assumption that the orthography is a kind of handicap that makes the acquisition of encoding skills relatively slower in Danish children than in Icelandic children even beyond the earliest phases of spelling acquisition.

Cross-linguistic comparisons of decoding and spelling development

If specific effects of orthographic depth are to be observed in cross-linguistic comparison, it is an important requirement that the comparisons are based on items that are as similar as possible across languages. Obviously, differences with respect to word structure or frequency could lead to differences in performance levels that have nothing to do with orthographic depth. The cross-linguistic matching of items has been quite convincing in several previous studies. For instance, Seymour, Aro and Erskine (2003) found that Danish and Scottish first graders read simple words and nonwords less accurately than first graders in European countries with more transparent orthographies, even though the items were highly similar across languages. This result supports the assumption that reading acquisition is more difficult in deep orthographies such as Danish and English.

Similar results have been obtained in comparisons of reading as well as spelling in English and German-speaking (Austrian) children aged 7 to 9 (Wimmer & Goswami, 1994; Wimmer & Landerl, 1997; Frith, Wimmer & Landerl, 1998). These authors exploited the linguistic relatedness of English and German and compared performances for words that had similar meanings as well as similar structures and frequencies. For instance, the English word *bread* was matched to the cognate German word *Brot*. In these studies the English participants were outperformed by the Austrians both on reading and spelling. Thus, even when items were very strictly matched, literacy development was found to be slower in the deep English orthography (cf. also Aro & Wimmer, 2003).

Another important requirement of studies that aim to isolate effects of orthographic depth is that the samples differ with respect to orthography, but not with respect to other factors that might influence literacy development. This ideal requirement is somewhat more difficult to meet. If linguistic groups are matched, for instance, on a math test (Ellis & Hooper, 2001), one can assume that they do not differ dramatically with respect to educational achievement in general. If they are matched on standardised national tests of reading age one can even assume that they are representative for their educational backgrounds (e.g. Frith *et al.*, 1998; Goswami *et al*., 2003). However, such matching procedures do not rule out the possibility that cross-linguistic differences are due to other factors than orthography. Literacy standards may differ across countries for reasons that have nothing to do with orthography. Different standards could reflect socio-economic or cultural differences, for instance, and, unfortunately, it can be very difficult to control such differences across countries.

The solution adopted in the present study was to match groups across languages on relatively simple spelling tasks, using cross-linguistically matched items to ensure comparability. For instance, we identified groups of Danish and Icelandic children who performed on the same level when they had to represent single consonant phonemes in onset position. Then we went on to compare their performances on more demanding items with consonant cluster onsets or medial doublets. For reasons explained below we suspected that the scores on these more demanding items would be selectively affected by orthographic depth. Socio-economic and cultural differences between the samples could still have effects on spelling performances, of course, but we would expect such effects to be general, rather than selective.

We now turn from methodology to a theoretical issue: Are the effects of orthographic depth found only when items differ with respect to the regularity of sound-letter correspondences? It is perhaps not so surprising to find an effect where sound-letter correspondences are more regular in one of the orthographies compared. In the English-Austrian comparison of spelling referred to above it was found that the vowel of German *Brot* was spelled correctly by 99% of the Austrians, while the vowel of the English cognate word *bread* was spelled correctly only by 60% of the English participants (Wimmer & Landerl, 1997). Clearly, the English vowel spelling was less regular than the German vowel spelling, because it involved a digraph, *ea*, and because a much more common spelling for English /e/ exists, viz. *e* as in *red* (Carney, 1994). It seems logical to assume that the difference in accuracy was due to the difference in regularity. This interpretation is supported by the fact that only 53% of the Austrian participants spelled the vowel correctly in the German word *Boot* "boat" where the less common digraph spelling *oo* is used for the same vowel as in *Brot*. This error

pattern suggests that spelling problems occur specifically when sound-letter correspondences are irregular.

It is also possible, however, that an effect of orthographic depth can be found even where sound-letter correspondences are similar across orthographies. A case in point is the correspondences between spoken and written consonant clusters in the Germanic languages. These correspondences are quite regular, not only in the more transparent orthographies, like German and Icelandic, but also in English and Danish. Interestingly, reductions of consonant clusters is a spelling error type that seems to persist for some time in English-speaking children (Snowling, 1994; Treiman, 1998). By contrast, observations of Austrian children's spelling suggest that this error type is almost non-existent in German-speaking children by the end of Grade 1 (Wimmer & Landerl, 1997). To explain this apparent difference, Wimmer and Landerl proposed that reading the transparent German orthography, paired with an instructional emphasis on sound-letter correspondences, help Austrian children to discover the segmental structure of consonant clusters.

This explanation may be correct, but strong evidence in support of it is wanting. The observation that Austrian first graders made few spelling errors on consonant clusters was based on just 13 clusters occurring in 11 words – e.g. *schwimmt* "swims" and *trägt* "carries". Word-specific spelling knowledge may have contributed to the good Austrian results, and it seems possible that problems with clusters would have shown if more items had been included. Furthermore, the assumption that the frequency of spelling errors on consonant clusters varies as a function of orthographic depth should be tested in a direct comparison. The study reported here is an attempt to fill this gap.

The spelling of consonant clusters has been the focus of one previous cross-linguistic study. Caravolas and Bruck (1993) compared the spelling of Czech and English-speaking (Canadian) first graders, and found that the English-speaking participants made a significantly greater number of errors on consonant clusters in nonwords than the Czech participants did. Both groups performed at ceiling when spelling singleton consonants. This result is in line with Wimmer and Landerl's reasoning, as Czech orthography, like German, is relatively transparent.

However, Caravolas and Bruck found that Czech and English also differ with respect to syllable structure. For instance, they observed that the frequency and variety of consonant clusters in onset position is greater in Czech. Thus, the superior spelling performance of the Czech participants may be due, at least in part, to a difference in familiarity with consonant clusters rather than to the difference in orthography. Indeed, these authors found that Czech pre-schoolers were better than English-speaking pre-schoolers at identifying initial phonemes of words with initial consonant clusters.

Using similar phoneme-identification tasks Birgisdóttir and Bryant (2000) found that Icelandic pre-schoolers, too, performed better than English-speaking pre-schoolers. As these authors found the variety and frequency of cluster onsets to be greater in Icelandic than in English, this result also points to a possible cross-linguistic effect of syllable structure (cf. also Seymour *et al*., 2003). Thus, if specific effects of orthographic depth are to be observed, cross-linguistic comparisons should ideally be based on languages that are structurally similar, but differ in orthographic depth. As we will argue below, Danish and Icelandic are languages well suited for such comparisons.

A Danish–Icelandic comparison of spelling

Our Danish–Icelandic comparison of spelling focused on the spelling of two specific phonological structures: consonant phonemes in word medial (intervocalic) position and consonant clusters in onset position.

The spelling of medial consonants was compared because sound-letter correspondences in Danish and Icelandic differ considerably. In both orthographies doublets are sometimes used to spell medial consonants, but the doubling rule is more complex in Danish: Danish doublets mark that a preceding stressed vowel is short – much like English medial doublets (e.g. *tapping* vs. *taping*). Icelandic doublets also follow short stressed vowels, but usually the doubling also marks that the medial consonant itself is phonetically long. This is not the case in Danish. Furthermore, the doubling rule is relatively more consistent in Icelandic. For instance, it applies to word final consonant, too, whereas in Danish it does not. Our expectation was that the more complex and inconsistent correspondences between consonant phonemes and doublets would lead to more doubling errors in the Danish group.

The spelling of consonant clusters was compared for the reasons already described above: Previous studies have suggested that errors on clusters are especially frequent in deep orthographies. Furthermore, consonant cluster spellings are equally regular in Danish and Icelandic and thus provide a possibility to test the hypothesis that error frequencies are associated with orthographic depth even when the regularity of sound-letter correspondences is controlled. We suspected that if the reasoning of Wimmer and Landerl (1997) was correct we would find more spelling errors on consonant clusters in the Danish group, because sound-letter correspondences are *generally* more irregular in Danish – although not in the specific case of consonant clusters!

Errors on consonant clusters and (especially) on consonant doublets are well-known spelling error types in both languages. However, to the best of our knowledge, the present study is the first empirical study in both countries that has explored the frequency of these error types in more detail.

Before we go on to report our study, we should like to demonstrate some crucial differences and similarities between Danish and Icelandic. The depth of an orthography

can be measured on at least two dimensions (Van den Bosch, Content, Daelemans & De Gelder, 1994). One is the consistency of sound-letter correspondences, that is, how easily spellings can be predicted from pronunciations and vice versa. Another is the frequency of deviations from the principle of "one sound per letter". Traditionally, most emphasis is put on the first dimension, but the second dimension of orthographic depth may have an impact on literacy development, too (Berndt, D'Autrechy & Reggia, 1994; Elbro, submitted). When graphemes do not always correspond to single letters, the segmentation of words into functional pronunciations units (graphemes) will tend to become more difficult – whether the resulting graphemes are consistent or not.

Danish orthography is deeper than Icelandic orthography both on the first and the second dimension. As for consistency, the difference is especially clear when the sound-letter correspondences for vowels are compared. Most Danish vowel phonemes correspond to more than one spelling, and vice versa, whereas most Icelandic vowel qualities correspond to just one spelling, and vice versa.

As for deviations from the "one sound per letter" principle we made a comparison based on the 500 most frequent words in each language (Bergenholtz, 1992; Pind, Magnússon & Briem, 1991). Deviations were defined relative to maximally distinct standard pronunciations. The comparison showed that deviations were far more common in Danish than in Icelandic (Table 1). Thus, the assumption also made in previous studies (Elley, 1992; Seymour *et al*., 2003) that Danish orthography is deeper than Icelandic orthography was confirmed. The frequency of consonant doublets, which may be said to represent just one sound (whether short or long), were not counted as deviations, as they were counted separately (cf. below). Diphthongs represented by a single vowel letter (cf. English words such as *find and most*) were not counted as deviations either, although they have a two-phase pronunciation. Such diphthongs were found in 94 Icelandic words, while they did not occur at all in the Danish word list. It should be noted, however, that Icelandic diphthongs "behave" like single vowel phonemes: Just like monophthongs they can be either short or long. Similarly, the two-phased pronunciation of the Danish /t/-phoneme – which sounds like /ts/ to many foreigners – was counted as a single sound.

The lists of frequent words were also used to confirm the structural similarity between Danish and Icelandic. As our study focused on consonant doublets in intervocalic position and consonant clusters in onset position, we compared the frequencies of these orthographic structures. The frequencies were roughly on the same level, both for types and tokens (Table 1). These similarities make it rather unlikely that cross-linguistic differences in spelling accuracy for these structures can be explained by differences in the amount of experience that participants have had with the same structures. We note that four-phoneme cluster onsets occur, although very rarely, in Icelandic – e.g. *skrjóður* "jalopy" – but not in Danish. A particular striking cross-linguistic similarity was that clusters with an initial *s* followed by a stop (*p*, *t* or *k*) accounted for approximately one-third of all cluster tokens in both languages.

Table 1. *Frequency counts for words that deviate from the one sound per letter principle, words with intervocalic doublets, and words with initial consonant clusters. Counts are based on the 500 most frequent words in each language*

	Danish	Icelandic
One sound per letter-deviations	190	34
Doublets (intervocalic)		
types	9	9
tokens	52	66
Clusters (word initial)		
types	22	19
tokens	67	54

METHOD

Participants

A total of 115 Danish and 77 Icelandic children participated in the study (Table 2 provides an overview). Participants came from socially mixed areas outside the centres of the capital cities, Copenhagen and Reykjavík. The Danish participants all came from one school, while the Icelandic participants came from two different schools.

There are several social and cultural differences between Copenhagen and Reykjavík. An obvious difference is that Copenhagen is much bigger than Reykjavík. Another difference is that children in Copenhagen are more often bilingual (cf. Table 2). Participants were asked if Danish/Icelandic was the first language that they spoke, and if they spoke this language with their family. They were classified as bilingual if they responded *no* to either or both of these questions. As we thought it impossible to select participants with similar cultural backgrounds across countries, we decided *not* to exclude the results of bilingual children on a priori grounds.

Testing was done in the middle of the school year, but earlier in Denmark than in Iceland. On average the Icelandic participants had attended school about two months longer than the Danish. However, because Icelandic children begin school at the age of 6, and Danish children at the age of 7, the Icelandic fourth graders were roughly of the same age as the Danish third graders.

Table 2. *Overview of participants by grade level and country*

	Number		Age (months)		Male (%)		Bilingual (%)	
Grade	DEN	ICE	DEN	ICE	DEN	ICE	DEN	ICE
2	19	–	102	–	53	–	21	–
3	33	39	114	106	39	41	24	8
4	34	38	126	115	47	50	26	0
6	29	–	151	–	24	–	34	–

Note: 5 participants who failed to complete the test are not included here (1 Icelandic third grader, 2 Danish second graders, and 2 Danish third graders).

Measures

All measures were based on nonword spelling. The spelling tasks were constructed in Danish and Icelandic versions which were as similar as possible, as demonstrated in the procedure section below. All nonwords had two syllables. Only spellings of the target phonemes were scored.

Simple onsets. This measure was based on 33 items with a single consonant phoneme in onset position. These consonant phonemes were divided equally between 11 different qualities which were the same in both versions of the test. The standard spellings for these phonemes were *b, d, g, f, s, m, n, l, r, v* or *j* – but phonemically plausible alternatives (such as *c* in stead of *s*) were accepted.

Consonant cluster onsets. This measure was based on 41 items with consonant clusters in onset position: 33 two-phoneme clusters and 8 three-phoneme clusters. The 41 clusters were divided between 16 different types which were the same in the two versions of the test. They are listed here with their standard spellings, but all phonemically plausible spellings were accepted: *sv*, *fr* and *fj* (4 items each); *sp, st, sk, sm, sn, fl* and *sl* (3 items each); *str* and *stj* (2 items each); and *spr*, *skr*, *spj* and *skj* (one item each).

Word initial sp st sk. Seventeen of the items with consonant cluster onsets had an unvoiced stop consonant following a word initial /s/. Both in Danish and Icelandic the standard spellings for such onsets are *sp*, *st* and *sk*, respectively. However, the spellings *sb*, *sd* and *sg* are phonemically plausible alternatives in both orthographies. The frequency of the standard spellings was scored as a percentage of all phonemically plausible spellings.

Unfamiliar clusters. This measure was based on 6 items with two-phoneme clusters that would not normally occur in word initial position. These unfamiliar clusters were: *mb, nt, ks, pf, vl* and *nj* (Danish version) or *rg* (Icelandic version). No other spellings were accepted. The onsets *mb*, *nt*, *ks* and *pf* do not occur in any word in any of the two languages, while *vl* occurs only in a Russian name (Vladimir) of low frequency in both languages. The onset *nj* which was used for Danish occurs in the famous Icelandic name Njál. It was replaced in the Icelandic version by *rg* which does not occur in any Icelandic word.

Doublets. This measure was based on 24 items with a consonant phoneme in intervocalic position. These were divided equally between two types: 12 consonants preceded by a short stressed vowel (in the Danish version this consonant was phonetically short, while in the Icelandic version it was long), and 12 consonants preceded by a long stressed vowel (in both versions this consonant was phonetically short). The consonants preceded by short vowels should be doubled both in Danish and Icelandic spelling. The consonants preceded by long vowels should *not* be doubled and were included to control the extent to which the use of doublets was over-generalized to consonants where doubling was not appropriate.

The consonant phonemes were divided equally between four different qualities, whether they should be doubled or not: /m n s/ and /f/ (Danish verison) or /r/ (Icelandic version). In the Danish version the spelling *nd* was accepted as an equivalent of the doublet *nn* because of the homophony of word pairs such as *ve**nn**er* "friends" and *ve**nd**er* "turns" (with a silent *d*).

Procedure

Participants were given a booklet with 12 preprinted nonwords on each page. In each nonword one or more letters had been replaced by an underscore. The nonwords were structurally very similar in the Danish and Icelandic versions of the test, as illustrated by the following examples (with IPA-transcriptions and correct spellings added):

Danish	*Icelandic*
__use /'duːsə/ (***d**use*)	__úla /'duːla/ (***d**úla*)
__yle /'sbryːlə/ (***spr**yle*)	__ila /'sbrɪːla/ (***spr**ila*)
ja__e /'jɑmə/ (*ja**mm**e*)	ja__a /'jamːa/ (*ja**mm**a*)
ba__e /'bæːmə/ (*ba**m**e*)	ba__a /'baːma/ (*ba**m**a*)

Participants were asked to fill in the missing letters as the words were dictated to them. A total of 104 items were used for the measures described above. In both versions of the test, however, additional items were included for other purposes than the present study (71 in the Danish version, and 112 in the Icelandic version).

The test was administered to one class at a time in the participants' own classroom. Items were presented and repeated twice by a test leader (not the same person in Copenhagen and Reykjavík) who had been instructed about how each item should be pronounced. Time was allowed until all participants were ready to proceed to the next item. Testing was always completed within two lessons. The whole test was administered on the same day in Copenhagen (with breaks). In Reykjavík it was administered over two days. Items were presented in a fixed random order. The six items used to measure performances with unfamiliar clusters were presented separately at the end of the test.

RESULTS

Table 3 shows the main results in the two countries by grade level. Performances on simple onsets were generally close to

Table 3. *Mean percentages correct (with SDs) by grade level and country on simple onsets, cluster onsets and doublets (percentages correct vs. over-generalized)*

					Doublets			
	Simple onsets (n = 33)		Cluster onsets (n = 41)		Correct (n = 12)		Over-generalized (n = 12)	
Grade	DEN	ICE	DEN	ICE	DEN	ICE	DEN	ICE
2	79 (23)	–	35 (31)	–	1 (3)	–	0.5 (2)	–
3	94 (5)	86 (21)	58 (33)	81 (23)	11 (26)	46 (42)	5 (15)	4 (12)
4	89 (15)	96 (5)	75 (25)	92 (12)	17 (26)	62 (38)	4 (9)	5 (14)
6	92 (9)	–	84 (16)	–	51 (43)	–	7 (13)	–

ceiling, and there was no clear-cut cross-linguistic difference on this measure. On cluster onsets and doublets, by contrast, the Icelandic participants performed much better than Danish participants at the same grade level. The Icelandic third graders even outperformed the Danish fourth graders, who were almost two years older, and they did nearly as well as the Danish sixth graders, who were almost four years older! There was hardly any difference across languages in the frequency of over-generalized doublets, suggesting that only word structures where doubling is required were affected by language.

A 3 × 2 × 2 analysis of variance with repeated measures was performed with simple onsets, cluster onsets and correctly used doublets as within-subject factors and grade level (third vs. fourth grade only) and language as between-subjects factors. There were significant main effects of spelling condition ($F(2, 280) = 297.452$, $p < 0.001$), as well as grade level ($F(1, 140) = 8.222$, $p < 0.01$), and country ($F(1, 140) = 38.783$, $p < 0.001$). The effect of spelling condition interacted with grade level ($F(2, 280) = 3.267$, $p < 0.05$) as well as country ($F(2, 280) = 34.354$, $p < 0.001$). This was because the effects were largest in the third graders and in the Danish participants.

Planned comparisons showed that both the contrasts between single and cluster onsets and between cluster onsets and doublets were significant ($F(1,140) = 85.662$, $p < 0.001$ and $F(1, 140) = 228.132$, $p < 0.001$, respectively). Both contrasts interacted with country ($F(1, 140) = 43.3$, $p < 0.001$ and $F(1, 140) = 12.079$, $p = 0.001$), but only the contrast between simple and cluster onsets interacted with grade level ($F(1, 140) = 13.877$, $p < 0.001$). For this contrast there was even a three-way (condition by grade by language) interaction $F(1, 140) = 10.670$, $p = 0.001$), because the contrast was very large in the Danish third graders, but quite modest in the Icelandic third graders who spelled cluster onsets nearly as well as they spelled simple onsets.

The general results point to the conclusion, then, that Danish third and fourth graders have much larger problems with cluster onsets and doublets than their Icelandic peers. It should not be overlooked, however, that the Icelandic participants had their spelling problems, too. Even the Icelandic fourth graders performed far below ceiling on doublets.

The relative problems of the Danish third and fourth graders could reflect that these participants were unusually poor spellers for their grade levels, but this did not appear to be the case. A 3 × 4 analysis of variance with repeated measures was performed on all Danish participants with simple onsets, cluster onsets and correctly used doublets as within-subject factors and grade level as the between-subjects factor. This analysis showed a significant overall effect of grade level ($F(3, 111) = 18.948$, $p < 0.001$). Importantly, post hoc tests showed that the third and fourth graders performed significantly above the second graders and significantly below the sixth graders (Scheffé, $p < 0.05$), although they did not differ significantly from each other. Thus, the spelling level of the Danish third and fourth graders did not appear abnormal.

The relatively poorer Danish results could not be ascribed to the larger proportion of bilingual participants either. When bilinguals were excluded, the Danish results were found to be similar to or *worse* than those reported in Table 3. For this reason the results of the bilingual participants were not excluded from the analyses.

We note, also, that the reliability of the measures was satisfactory. Reliability coefficients (alpha) were computed for all third and fourth graders in each country ($N = 67$ in the Danish group, and $N = 77$ in the Icelandic group). For these analyses items were scored as correct vs. all other response types. Danish coefficients are reported first, Icelandic last. Simple onsets: 0.84 and 0.93. Cluster onsets: 0.97 and 0.94. Correctly used doublets: 0.93 and 0.96.

Still, the specific cross-linguistic differences on cluster onsets and doublets in Table 3 may only reflect a general difference in spelling skills between the Danish and Icelandic participants. Because of the ceiling effect on simple onsets, it remains possible that this measure was not sensitive enough to capture a genuine cross-linguistic difference in spelling skills at this level. Thus, the differences on cluster onsets and doublets may not be specific to these particular measures. It also remains possible that the Danish sample happened to include a larger proportion of poor spellers than the Icelandic sample.

To counter such doubts we next report some more restricted cross-linguistic comparisons. For these comparisons participants from the third and fourth grades were matched on their performances on simple onsets or cluster onsets. To increase the comparability of the samples, participants were excluded from the comparisons if they performed at ceiling, and also if their scores were low. Table 4 shows the results for participants matched on their spelling

Table 4. *Participants matched on their spelling of simple onsets. Mean percentages correct (with SDs) on simple onsets, cluster onsets and doublets (percentages correct vs. over-generalized)*

				Doublets	
Country	*N*	Simple onsets ($n = 33$)	Cluster onsets ($n = 41$)	Correct ($n = 12$)	Over-generalized ($n = 12$)
DEN	28	91 (3)	67 (27)	8 (17)	4 (12)
ICE	28	90 (3)	87 (15)	51 (39)	4 (13)

Table 5. *Participants matched on their spelling of cluster onsets. Mean percentages correct (with SDs) on cluster onsets, doublets (percentages correct vs. over-generalised), word initial* sp st sk*, and unfamiliar clusters.*

			Doublets			
Country	*N*	Cluster onsets (n = 41)	Correct (n = 12)	Over-generalised (n = 12)	*sp st sk* (n = 17)	Unfamiliar clusters (n = 6)
DEN	31	89 (7)	20 (29)	7 (15)	89 (17)	34 (24)
ICE	29	89 (4)	41 (35)	6 (14)	82 (26)	59 (30)

of simple onsets. Only Danish spellers who made from 2 to 5 errors and Icelandic spellers who made from 2 to 7 errors on the 33 simple onset items were included in this comparison.

This more restricted comparison showed cross-linguistic differences similar to those in Table 3. There was a strong effect of spelling condition ($F(2, 108) = 156.309$, $p < 0.001$) as well as language ($F(1, 54) = 27.875$, $p < 0.001$). The effect of spelling condition interacted significantly with language ($F(2, 108) = 18.220$, $p < 0.001$), as the effect of spelling condition was much stronger in the Danish sample. Planned comparisons showed, again, that both the contrasts between single and cluster onsets and between cluster onsets and doublets were significant ($F(1, 54) = 25.578$, $p < 0.001$ and $F(1, 54) = 134.033$, $p < 0.001$, respectively), and, again, both contrasts interacted significantly with language ($F(1, 54) = 14.996$, $p < 0.001$ and $F(1, 54) = 7.596$, $p < 0.01$). Again, the frequency of over-generalizations was remarkably similar across languages.

Thus, this more restricted comparison confirms the impression that, relative to Icelandic children, Danish children have specific spelling problems with consonant clusters and consonant doublets.

We also made a comparison of groups matched on their performances with cluster onsets to check if a cross-linguistic difference would still be found for consonant doublets if the match was based on a more difficult spelling task. A further purpose was to compare performances on the more demanding measures of consonant cluster spelling: the use of standard spellings for stops after *s* (i.e., *sp st sk*) and the spelling of unfamiliar cluster onsets such as *mb*. Again, the very best and the very poorest spellers were excluded. Only Danish participants who made from 1 to 10 errors and Icelandic participants who made from 3 to 10 errors on the 41 cluster items entered the comparison.

The performances on onset clusters (the matched variable) were compared to correctly used doublets, the use of *sp st sk*, and unfamiliar clusters in three 2×2 analyses of variance with repeated measures. For cluster onsets vs. doublets there were significant effects of spelling condition ($F(1, 58) = 215.565$, $p < 0.001$) as well as language ($F(1, 58) = 5.610$, $p < 0.05$), and as expected the interaction was significant because the effect of spelling condition was stronger in the Danish participants ($F(1, 58) = 7.389$, $p < 0.01$). Similarly, for (familiar) cluster onsets vs. unfamiliar clusters there were significant effects of spelling condition ($F(1, 58) = 166.066$, $p < 0.001$) and language ($F(1, 58) = 9.586$, $p < 0.01$), and, again, the larger Danish effect of spelling condition led to a significant interaction ($F(1, 58) = 14.784$, $p < 0.001$). For cluster onsets vs. *sp st sk*, however, no significant effects were found.

These comparisons provide even stronger support for the conclusion that the use of doublets is a greater spelling problem in Danish than in Icelandic. They also gave further support to the impression that the representation of consonant clusters is affected by the orthographic factor. Although the groups performed on the same level with familiar consonant clusters, there was a very clear difference on unfamiliar clusters. This latter measure was based on only six items, but the difference does suggest that the Danish spellers relied relatively more than their Icelandic peers on previous experience with the clusters presented, and relatively less on an immediate segmental analysis. Finally, the fact that the Danish and Icelandic groups did not differ much when it came to over-generalizations of doublets and to the use of the standard spellings *sp, st* and *sk* suggests that the spelling problems of the Danish group were indeed confined to specific linguistic structures.

DISCUSSION

The present study was an attempt to isolate the effects of orthographic depth on the acquisition of spelling skills in Danish and Icelandic. The results support the conclusion that orthographic depth can be described as a kind of handicap for children learning to spell.

Children learning to spell in the deep Danish orthography were outperformed by children learning to spell in the more transparent Icelandic orthography when we compared the use of doublets for medial consonants – in spite of the fact that both groups performed on the same level on simple spelling tasks. Because the doubling rule is simpler and more consistent in Icelandic, this result was expected. In fact, the cross-linguistic difference can be seen as a parallel to the finding that the unpredictable vowel spellings of English orthography are a greater spelling problem than the more predictable vowel spellings of German orthography (Wimmer & Landerl, 1997).

The Danish participants were also outperformed when we looked at the accuracy with which familiar and unfamiliar

consonant clusters were spelled. Apparently, Danish children are less efficient than their Icelandic peers when they have to use writing as a sound script. This was perhaps more surprising, as the sound-letter correspondences of consonant clusters are perfectly regular in both orthographies. Such a difference would be expected, however, if Wimmer and Landerl (1997) were correct in claiming that learning to read a transparent orthography provides a better training of phoneme segmentation skills than learning to read a deep orthography. The idea that reading a deep orthography such as Danish makes it relatively harder to discover the phonemic structure of words seems to go well with the fact that we found deviations from the principle of "one sound per letter" to be very common in Danish words.

We must point out that differences found between matched groups can sometimes be misleading. If the match had been based on the poorest Icelandic and the best Danish spellers, measurement error would be expected to yield an overestimate of the performance level for Danish spellers, but an underestimate of the performance level for Icelandic spellers, and differences found on other measures could simply be a regression toward the "true" spelling ability of the groups compared (Cook & Campbell 1979). It is not likely, however, that the differences found between matched groups in the present study can be explained as a regression effect, as all participants with extreme scores were excluded from the matched comparisons. Furthermore, the reliability of the measures was found to be high.

The study also supported Wimmer and Landerl's claim that the spelling of consonant clusters is only a marginal problem in the acquisition of spelling in a transparent orthography. As our measures were based on more items, they were presumably more sensitive than measures in previous studies. Even so, the Icelandic children were found to represent consonant clusters rather accurately, much like the Czech children studied by Caravolas and Bruck (1993) and the Austrian children studied by Wimmer and Landerl (1997).

The Icelandic participants did show problems on the use of consonant doublets, however. This suggests that the representation of phonemic quantity may be a general problem in learning to spell, even when the marking of quantity is quite straightforward. Problems with doublets are known also from Greenlandic and Finnish, where consonant doublets, as in Icelandic, correspond to phonetically long consonants (Jacobsen, 1994; Lyytinen, Leinonen, Nikula, Aro & Leiwo, 1995).

As pointed out earlier, the fact that the Danish and Icelandic participants performed on the same level on simple spelling tasks makes it less likely that there was a *general* difference in spelling skills between the groups. However, this still does not rule out the possibility that other factors than orthographic depth could lead to the *specific* differences found on onset clusters and doublets. We will discuss two such factors.

First, just as a transparent orthography may favour the development of basic decoding and encoding skills, so may a phonics-based instructional approach (Wimmer & Landerl, 1997; Bruck, Treiman, Caravolas, Genesee & Cassar 1998; Landerl, 2000). We do not know whether the Icelandic teachers were more phonics-oriented than the Danish teachers, but it is certainly possible. However, the very fact that the sound-letter correspondences of Danish orthography appear so irregular may be one of the most important reasons if Danish teachers tend not to adopt a purely phonics-oriented approach to literacy instruction. It should also be pointed out that cross-linguistic differences have been found between English and German children's decoding skills even when the instructional approach was controlled (Landerl, 2000).

Second, specific differences in the degree of attention paid by teachers to errors on consonant clusters and doublets may, of course, lead to specific differences in spelling performances. To find out whether Danish teachers have a more tolerant attitude to certain spelling errors than their Icelandic colleagues, we distributed a questionnaire to all teachers at the participating Danish school and at a neighbouring school. The teachers were asked to indicate the grade level at which normal pupils should be able to detect and correct various examples of spelling errors, including errors on consonant clusters and doublets. According to the respondents (N = 12) normal Danish children should be able to correct cluster errors by the end of the third grade and doubling errors by the end of the fourth grade. Unfortunately, however, only 20% of the Danish teachers responded to the questionnaire. Because of this low response rate a planned Icelandic parallel was not implemented. Thus, we can only note the possibility that differences in attitudes towards specific spelling error types affected the results. Nevertheless, we believe that assessments of teacher attitudes may contribute interesting new data to future cross-linguistic comparisons. It should be added, though, that since the doubling rule is more complex in Danish than in Icelandic, it seems reasonable to suspect that Danish teachers have a more tolerant attitude to doubling errors.

What are the instructional implications of the cross-linguistic differences observed in the present study? If the acquisition of phonemic encoding skills is more difficult in a deep orthography, then instructional practices should definitely take account of this fact. However, to do this, it is important to understand the nature of the problem.

Some may conclude that what Danish spellers need to learn must be something very different from what Icelandic spellers need to learn, given that the relations between spoken and written language in Danish differ so much from those in Icelandic. This is not the conclusion we would draw. As shown, there are some clear differences between Danish and Icelandic orthography – e.g. in the spelling of vowels and intervocalic consonant phonemes and in the frequency of silent letters. However, there are also great similarities, e.g.

in the spelling of consonant clusters. At least for consonant clusters, then, what spellers need to learn appears to be almost the same in the two languages.

It should be noted, also, that although deviations from the principle of "one sound per letter" are frequent in Danish, we did find that more than 60% of the 500 most frequent Danish words had a perfect match between the number of sounds and letters. Furthermore, it is well-known that Danish pre-schoolers' phonemic awareness is a strong predictor of later development of literacy skills (Lundberg, Frost & Petersen, 1988; Petersen & Elbro, 1999). This would hardly be the case if knowledge of sound-letter correspondences did not play an important role for this development, even in Danish.

Thus, rather than teaching something completely different from basic sound-letter correspondences, the challenge to teachers of Danish and other deep orthographies may be to pay special attention to children's knowledge of these correspondences. One way of helping could be to reduce the number of irregular words in early reading and spelling instruction (Borstrøm, Petersen & Elbro, 1999).

Danish children do have to learn about the irregularities of their orthography, of course. However, there is reason to believe that the irregularities are easier to cope with when a sound knowledge of regular sound-letter correspondences has been achieved (Caravolas, Hulme & Snowling, 2001). Assisting children in developing such knowledge may be an instructional challenge of special importance in deep orthographies.

The Danish part of this study was made possible by a grant from the Danish Research Council for the Humanities (#9602075) to Professor Carsten Elbro, Centre for Reading Research, University of Copenhagen. The Icelandic part of this study was supported by a grant from the Research Fund of the Iceland University of Education. We would like to thank the following people for practical assistance: Tinne Ferslev, Maja Dyrby and Morten Bruun Petersen (Denmark); Rannveig G. Lund, Ásta Lárusdóttir and Steinunn Torfadóttir (Iceland). Special thanks to all the pupils and teachers who participated in the study.

REFERENCES

Allerup, P., Mejding, J. & Zeuner, L. (2001). *Færdigheder i læsning og matematik: udviklingstræk omkring årtusindskiftet* [Reading and math abilities. Developmental trends around the turn of the millennium]. Copenhagen: AKF, F2000 sekretariatet.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, *24*, 621–635.

Bergenholtz, H. (1992). *Dansk frekvensordbog* [Danish frequency dictionary]. Copenhagen: G.E.C. Gad.

Berndt, R. S., D'Autrechy, C. L. & Reggia, J. A. (1994). Functional pronunciation units in English words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 977–991.

Birgisdóttir, F. & Bryant P. (2000). *The analysis of word-initial and word-final clusters by Icelandic and English children.* Poster presented at the SSSR-conference in Stockholm, July 2000.

Borstrøm, I., Petersen, D. K. & Elbro, C. (1999). *Hvordan kommer børn bedst i gang med at læse? En undersøgelse af læsebogens betydning for den første læseudvikling* [How do children best begin to read? A study of the primer's impact on reading development]. Copenhagen: Center for Læseforskning.

Bruck, M., Treiman, R., Caravolas, M., Genesee, F. & Cassar, M. (1998). Spelling skills of children in whole language and phonics classrooms. *Applied Psycholinguistics*, *19*, 669–684.

Caravolas, M. & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology*, *55*, 1–30.

Caravolas, M., Hulme, C. & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, *45*, 751–774.

Carney, E. (1994). *A survey of English spelling*. London: Routledge.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation. Design & analysis issues for field settings*. Chicago: Rand McNally College Publishing Company.

Elbro, C. (submitted). *Literacy acquisition in Danish: a deep orthography in cross-linguistic light.* Manuscript submitted for publication. Copenhagen: University of Copenhagen.

Elley, W. B. (1992). *How in the world do students read?* Hamburg: The International Association for the Evaluation of Educational Achievement.

Ellis, N. C. & Hooper, M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics*, *22*, 571–599.

Frith, U., Wimmer, H. & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, *2*, 31–54.

Frost, R. (1992). Orthography and phonology. The psychological reality of orthographic depth. In P. Downing, S. D. Lima & M. Noonan (Eds.), *The linguistics of literacy*. Amsterdam: John Benjamins.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics*, *24*, 235–247.

Jacobsen, B. (1994). "The mystery of quantity": Quantity as a specific spelling problem, and probably a reading problem. *Acta Linguistica Hafniensia*, *27*, 235–259.

Juul, H. (2001). Fra analfabet til avanceret ortografibruger [From illiterate to advanced orthography user]. *Danske Studier 2001*, 85–113.

Katz, L. & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning*. Amsterdam: North-Holland.

Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of Psychology of Education*, *XV*, 239–257.

Lundberg, I. (1997). Does language make any difference in reading acquisition? In Å. Olofsson & S. Strömqvist (Eds.), *Cross-linguistic studies of dyslexia and early language development*. Luxembourg: Office for Official Publications of the European Communities.

Lundberg, I., Frost, J. & Petersen, O.-P. (1988). Effects of an extensive program for stimulating phonological awareness in pre-school children. *Reading Research Quarterly XXIII*, 263–284.

Lyytinen, H., Leinonen, S., Nikula, M., Aro, M. & Leiwo, M. (1995). In search of the core features of dyslexia: Observations concerning dyslexia in the highly orthographically regular Finnish language. In V. W. Berninger (Ed), *The varieties of orthographic knowledge II: Relationships to phonology, reading, and writing*. Dordrecht: Kluwer Academic Publishers.

Petersen, D. K. & Elbro, C. (1999). Pre-school prediction of dyslexia: A longitudinal study with children of dyslexic parents. In

T. Nunes (Ed), *Learning to read: An integrated view from research and practice*. Dordrecht: Kluwer Academic Publishers.

Pind, J., Magnússon, F. & Briem, S. (1991). *Íslensk orðtíðnibók* [Icelandic frequency dictionary]. Reykjavík: Orðabók Háskólans.

Seymour, P. H. K., Aro, M. & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Snowling, M. J. (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*. Chichester: John Wiley & Sons.

Sommer, M., Lau, J. & Mejding, J. (1996). *Nordlæs – en nordisk undersøgelse af læsefærdigheder i 1.-3. klasse* [Nordlæs – a Nordic study of reading skills in Grades 1 to 3]. Copenhagen: Danmarks Pædagogiske Institut.

Treiman, R. (1998). Beginning to spell in English. In C. Hulme & R. M. Joshi (Eds.), *Reading and Spelling: Development and Disorders*. Mahwah, New Jersey, Lawrence Erlbaum Associates, Publishers.

UNICEF (2002). *A league table of educational disadvantage in rich nations*. Florence: Innocenti Research Centre.

Van den Bosch, A., Content, A., Daelemans, W. & De Gelder, B. (1994). Measuring the complexity of writing systems. *Journal of Quantitative Linguistics*, *1*, 178–188.

Wimmer, H. & Goswami, U. (1994). The influence of orthographic consistency on reading development: word recognition in English and German children. *Cognition*, *51*, 91–103.

Wimmer, H. & Landerl, K. (1997). How learning to spell German differs from learning to spell English. In C. A. Perfetti, L. Rieben & M. Fayol (Eds.), *Learning to Spell. Research, Theory, and Practice across Languages*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science*, *12*, 379–384.