Becoming literate in different languages: similar problems, different solutions

Johannes C. Ziegler¹ and Usha Goswami²

¹. Laboratoire de Psychologie Cognitive, CNRS and University of Provence, France
². Faculty of Education, University of Cambridge, UK


Abstract

The teaching of reading in different languages should be informed by an effective evidence base. Although most children will eventually become competent, indeed skilled, readers of their languages, the pre-reading (e.g. phonological awareness) and language skills that they bring to school may differ in systematic ways for different language environments. A thorough understanding of potential differences is required if literacy teaching is to be optimized in different languages. Here we propose a theoretical framework based on a psycholinguistic grain size approach to guide the collection of evidence in different countries. We argue that the development of reading depends on children's phonological awareness in all languages studied to date. However, we propose that because languages vary in the consistency with which phonology is represented in orthography, there are developmental differences in the grain size of lexical representations, and accompanying differences in developmental reading strategies across orthographies.

Introduction

Reading is about gaining access to meaning from printed symbols. To access meaning from print, the child must learn the code used by their culture for representing speech by a series of visual symbols. The first steps in becoming literate, therefore, require acquisition of the system for mapping distinctive visual symbols onto units of sound (phonology). This mapping process is called phonological recoding (Share, 1995). Mastery of this process allows the child to access the thousands of words that are present in their spoken lexicons prior to reading, and also to recode words that they have heard but never seen before. Phonological recoding works well as a self-teaching device, because the relationship between symbol and sound is systematic in most languages (e.g. the symbol ‘D’ is always pronounced /d/ at the beginning of a word). In contrast, mapping visual symbols directly onto units of meaning, as would be required by some sort of visual or ‘logographic’ learning, is difficult because the relationship between symbol and meaning is arbitrary. That is, knowing that a word starts with the letter ‘D’ tells the child nothing about its meaning.

It has become quite clear over recent years that visual learning does not represent a viable alternative to phonological recoding. Share (1995) compares visual learning to memorizing large slabs of a telephone directory:

Like printed letter strings, telephone numbers contain a small set of symbols . . . Unless all numbers are dialed correctly and in the right order the connection will fail . . . Unfortunately, there are no systematic or predictable relationships between these strings and their corresponding entries; so each of the many thousands of such associations must be painstakingly committed to memory. There may exist a few rare individuals who are capable of memorizing entire telephone directories, but for the normal child about to learn to read, the absurdity of this task should be obvious. (Share, 1995, p. 159)

Although phonological recoding is a much more efficient strategy than logographic learning, it nevertheless has a few problems of its own. The biggest problem has to do with inconsistency in the symbol-to-sound
mapping (Ziegler, Stone & Jacobs, 1997). In some orthographies, one letter or letter cluster can have multiple pronunciations (e.g. English, Danish), whereas in others it is always pronounced in the same way (e.g. Greek, Italian, Spanish). Similarly, in some orthographies, a phoneme can have multiple spellings (e.g. English, French, Hebrew), whereas in others it is almost always spelled the same way (e.g. Italian). English is exceptionally inconsistent because it exhibits a high degree of inconsistency in both directions (i.e. for reading and spelling). This variation across languages makes it likely that there will be differences in reading development across languages (and probably also in spelling development). It is relatively easy to learn about phonemes if one letter consistently maps onto one and the same phoneme, or if one phoneme consistently maps to one and the same letter. It is relatively difficult to learn about phonemes if a letter can be pronounced in multiple ways (e.g. the letter 'A' in English maps onto a different phoneme in the highly familiar words ‘cat’, ‘was’, ‘saw’, ‘made’ and ‘car’).

Reading development across languages

The most ambitious cross-language reading comparison to date has been conducted by the European Concerted Action on Learning Disorders as a Barrier to Human Development. Participating scientists from 14 European Community countries developed a matched set of items of simple real words and nonwords. These items were then given to children from each country during their first year of reading instruction (for details see Seymour, Aro & Erskine, 2003). Thus, children varied in age, but were equated for degree of reading instruction across orthography. Although method of reading instruction itself could not be equated exactly, schools were chosen so that all children were experiencing phoneme-level ‘phonics’ teaching. The results are summarized in Table 1.

The most striking finding from the study was that the children who were acquiring reading in orthographically consistent languages (Greek, Finnish, German, Italian, Spanish) were close to ceiling in both word and nonword reading by the middle of first grade. In contrast, English-speaking children lagged behind the French-speaking children by about 27% on nonword reading and 24% on word reading. Similar findings were obtained comparing beginning reading in English and Welsh. In parts of North Wales, English and Welsh are spoken and read side by side. In contrast to English, however, the writing system of Welsh is highly consistent. Parents choose whether they want their child to attend English or Welsh schooling. These schools serve the same geographical catchment area, are administered by the same local education authorities, and follow similar curricula and teaching approaches. The only real difference is the language of instruction. Yet, the results showed that the Welsh-speaking children could read well over twice as many words as the English-speaking children after the same amount of reading instruction.

Why is reading English so much more difficult?

One potential problem with these cross-language comparisons are socio-cultural differences across languages. For example, there may be differences in school systems, curricula, teaching methods and demographic distributions. This problem has begun to be addressed, however. Bruck, Genesee and Caravolas (1997) followed a group of English- and French-speaking children who were from the same area in Canada. They investigated word and nonword reading at the end of grade 1, using high-frequency regular monosyllabic words and nonwords. The results showed that the English-speaking children lagged behind the French-speaking children by about 27% on nonword reading and 24% on word reading. Similar findings were obtained comparing beginning reading in English and Welsh. In parts of North Wales, English and Welsh are spoken and read side by side. In contrast to English, however, the writing system of Welsh is highly consistent. Parents choose whether they want their child to attend English or Welsh schooling. These schools serve the same geographical catchment area, are administered by the same local education authorities, and follow similar curricula and teaching approaches. The only real difference is the language of instruction. Yet, the results showed that the Welsh-speaking children could read well over twice as many words as the English-speaking children after the same amount of reading instruction.

Table 1  Data (% correct) from the large-scale study of reading skills at the end of grade 1 in 14 European languages (adapted from Seymour, Aro & Erskine, 2003)

<table>
<thead>
<tr>
<th>Language</th>
<th>Familiar real words</th>
<th>Pseudowords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greek</td>
<td>98</td>
<td>92</td>
</tr>
<tr>
<td>Finnish</td>
<td>98</td>
<td>95</td>
</tr>
<tr>
<td>German</td>
<td>98</td>
<td>94</td>
</tr>
<tr>
<td>Austrian German</td>
<td>97</td>
<td>92</td>
</tr>
<tr>
<td>Italian</td>
<td>95</td>
<td>89</td>
</tr>
<tr>
<td>Spanish</td>
<td>95</td>
<td>89</td>
</tr>
<tr>
<td>Swedish</td>
<td>95</td>
<td>88</td>
</tr>
<tr>
<td>Dutch</td>
<td>95</td>
<td>82</td>
</tr>
<tr>
<td>Icelandic</td>
<td>94</td>
<td>86</td>
</tr>
<tr>
<td>Norwegian</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>French</td>
<td>79</td>
<td>85</td>
</tr>
<tr>
<td>Portuguese</td>
<td>73</td>
<td>77</td>
</tr>
<tr>
<td>Danish</td>
<td>71</td>
<td>54</td>
</tr>
<tr>
<td>Scottish English</td>
<td>34</td>
<td>29</td>
</tr>
</tbody>
</table>

© 2006 The Authors. Journal compilation © 2006 Blackwell Publishing Ltd.
inconsistent with respect to small reading units (letters or letter clusters corresponding to single phonemes). English is less inconsistent with respect to larger reading units, such as rimes or syllables (Treiman, Mullenix, Bijeljac-Babic & Richmond-Welty, 1995). This makes it likely that English children may be developing recoding strategies at more than one grain size. There are more orthographic units to learn when the grain size is big than when the grain size is small. For instance, in order to decode the most frequent 3000 monosyllabic English words at the level of the rime, a child needs to learn mappings between approximately 600 different orthographic patterns and 400 phonological rimes, far more than would be needed if the child could simply learn how to map 26 letters onto 26 phonemes. But relying solely on grapheme–phoneme correspondences leads to inefficient recoding of English. In contrast, young learners of relatively consistent languages can focus exclusively at the ‘small’ psycholinguistic grain size of the phoneme without making many reading errors. Consistent feedback received in terms of achieving correct pronunciations will further reinforce the acquisition process.

Psycholinguistic grain size theory (Ziegler & Goswami, 2005) takes these factors into account. The theory suggests that the dramatic differences in reading accuracy and reading speed found across orthographies reflect fundamental differences in the nature of the phonological recoding and reading strategies that are developing in response to the orthography. Children who are learning to read more orthographically consistent languages, such as Greek, German, Spanish or Italian, rely heavily on grapheme–phoneme recoding strategies because grapheme–phoneme correspondences are relatively consistent. Children who are learning to read less orthographically consistent languages, like English, cannot use smaller grain sizes as easily, because inconsistency is much higher for smaller grapheme units than for larger units like rimes. As a consequence, English-speaking children need to use a variety of recoding strategies, supplementing grapheme–phoneme conversion strategies with the recognition of letter patterns for rimes and attempts at whole word recognition.

Recent studies have indeed shown that inconsistent orthographies like English appear to push readers into developing both ‘small unit’ and ‘large unit’ recoding strategies in parallel (e.g. Brown & Deavers, 1999). To study the flexible and adaptive use of different grain sizes in reading English, Goswami and colleagues (Goswami, Ziegler, Dalton & Schneider, 2003) asked German and English children to read aloud several lists of nonwords. One list contained familiar orthographic patterns at a larger grain size (‘large unit’ nonwords, like daik and murn). Another list contained only unfamiliar large unit patterns (‘small unit’ nonwords, like dake and make). The prediction was that if a list contained only ‘large unit’ nonwords, then the exclusive application of a large grain size strategy should be very successful. In contrast, if the list contained only ‘small unit’ nonwords, then recoding should be most successful if an exclusively small grain size strategy was applied. The critical prediction was for a third ‘mixed’ list. If both types of nonwords were mixed within a particular list (e.g. daik, murn), continual switching between ‘small unit’ and ‘large unit’ processing may be required, incurring a switching cost. Consistent with this prediction, the results showed switching costs for the English children but not for the Germans. Blocking word lists by grain size apparently helped the English readers to focus at a single grain size, which increased recoding accuracy particularly for large unit items (like dake; here the child can use rhyme analogies to make, cake, bake, etc.). German readers did not show these ‘blocking’ effects. The absence of a blocking effect for the German readers was taken as evidence that they already relied exclusively on processing at the small grain size level.

Similar conclusions come from studies on reading aloud ‘sound-alike’ nonwords, so-called pseudohomophones (e.g. ‘faik’). It has been shown that English children show much stronger influences from whole word phonology than German children of the same age (Goswami, Ziegler, Dalton & Schneider, 2001). That is, English children showed a significant advantage in naming pseudohomophones in comparison to orthographic control nonwords (e.g. ‘faik’ read better than ‘daik’), whereas German children did not. This suggests that English children were more affected by whole word phonology when reading nonwords than German children. German children decoded nonwords that did not sound like real words as efficiently as nonwords that did sound like real words, resulting in an absence of the pseudohomophone effect in naming.

The important role of teaching

In alphabetic orthographies, the grain size problem is tackled by the teacher, who typically begins to teach reading from the single letter. The child is taught letter–sound correspondences, and hence learns about phonemes. As a result, phonological representations for words, which typically represent syllable and onset-rime information prior to reading, are rapidly augmented with phoneme-level information. Experience with written language changes the nature of phonological representations. In particular, it boosts phoneme awareness,
which, in turn, becomes the strongest predictor of successful reading, a reciprocal relationship (Perfetti, Beck, Bell & Hughes, 1987; Rayner, Foorman, Perfetti, Pesetsky & Seidenberg, 2001). Small grain size teaching works well in a language with consistent letter–sound correspondences, such as Italian. However, this teaching method works less well in a language with less consistent letter–sound correspondences, such as English.

One approach to this teaching problem in English-speaking countries has been to begin the teaching of reading at younger and younger ages, and to focus more and more at the phoneme level. In England itself, this approach is formalized in the National Literacy Strategy (DfEE, 1998), which requires the direct teaching of reading from the age of 5 years beginning with a phoneme-based strategy. Nevertheless, English children learn to read more slowly than children from other countries, who may not begin formal instruction until the age of 7 or even 8 years. For example, Finnish children begin school at 7, and are reading with 90% accuracy by approximately the tenth week in school (e.g. Seymour et al., 2003). English children who begin school at 4 or 5 years of age are still struggling to reach 90% accuracy (e.g. for non-word recoding) by age 9 or 10 (Goswami et al., 1998).

The slower average rate of learning to read in English does not seem to occur because of variations in teaching method across different countries. Rather, it seems due to the relatively low orthographic consistency of English. This was demonstrated for example in the English/Welsh comparison described earlier (Ellis & Hooper, 2001). Converging data come from Landerl (2000). She compared English children who were being taught to read by a ‘standard’ mixed method of phonics and whole word recognition with English children following a special phonics program that focused almost exclusively on letter–sound correspondences (this was a between-school comparison conducted before the advent of the National Literacy Strategy). She reported that the first grade English phonics children made almost as many errors on a nonword reading task (43%) as the first grade English standard children (50%, a non-significant difference) compared to 12% errors for a matched German sample. At second grade, a similar pattern was found (English standard = 29% errors, English phonics = 23% errors, German children = 13% errors). It was only by third grade that the English phonics children (7% errors) were comparable to the German children (14% errors), and by fourth grade that the English standard children reached a ‘German’ level of grapheme–phoneme recoding skill (12% errors compared to 11% for the Germans).

The teaching methods characteristic of early literacy instruction in Turkish provide even more striking evidence for the relatively minor impact of different phonic teaching regimes (at least, for children with normal phonological skills). Literacy instruction in Turkish elementary schools does not capitalize particularly on the high degree of orthographic transparency. In first grade, children are given sentences to memorize. They only receive instruction on individual components such as words, syllables or letters following successful rote memorization. Despite this ‘large to small’ grain size method, children usually get their ‘red ribbons’ symbolizing good decoding skills by December or January of the first school year (see Durgunoglu & Oney, 2002). In other countries with highly consistent orthographies, such as Greece and Hungary, some geographical areas have adopted the ‘whole language’ method of initial reading instruction (the whole language method has a meaning-based rather than a code-based emphasis; reading is characterized as a ‘psycholinguistic guessing game’ bootstrapped by meaning). It is not yet clear what the effects on initial reading acquisition will be. However, judging from the data in Table 1, they are likely to be quite minor for children with good phonological skills. The effects on children with poor phonological skills are likely to be rather greater.

A different approach to the question of the optimal grain sizes for teaching has been to begin instruction with correspondences for the larger units that are readily available in the phonological domain, such as rimes or syllables. As mentioned above, this approach requires the child to learn a much larger number of fairly complex letter combinations. Nevertheless, such a ‘large unit’ approach to teaching appears to lead to broadly similar progress in reading English as ‘small unit’ approaches. Recently, Walton and his colleagues (Walton, Bowden, Kurtz & Angus, 2001a; Walton & Walton, 2002; Walton, Walton & Felton, 2001b) compared the effectiveness of ‘large unit’ and ‘small unit’ approaches to the initial teaching of reading to children in Canada. In these studies, they compared the effects of teaching beginning readers to read by using a ‘rhyme analogy’ strategy (‘beak’ – ‘peak’, see Goswami, 1986) with the effects of teaching beginners to read by using a grapheme–phoneme recoding strategy. All the children were prereaders. Reading ability was assessed following 3 months of training, and four different kinds of words were used to assess different skills (analogy – irregular patterns [fight-sight], analogy – regular patterns [bed-teed], letter recoding [bat-bet] and nonwords [hibi]). Walton et al. found that both training groups showed broadly equal reading acquisition gains immediately following training. However, whereas the rhyme analogy group could also read new words requiring letter-recoding skills, the letter-recoding group could not read new words requiring rhyme analogy skills. When reading acquisition was
Developmental footprints on skilled reading

Do we need to consider developmental differences for our understanding of skilled reading? One could argue that the final product of reading development is the same in different languages despite different developmental trajectories. But is this the case? To address this issue, German and English skilled readers were asked to read aloud words that had highly similar spelling, sound and meaning in both languages (e.g. zoo in English versus Zoo in German; Ziegler, Perry, Jacobs & Braun, 2001). These words were not loan words in the respective languages, but genuine German and English words. Orthographic rime effects were used as a marker for ‘large unit’ processing and word length effects as a marker for ‘small unit’ processing. It was expected that German readers would show stronger length effects for the same items compared to English readers, because their basic processing unit is small. In contrast, English readers were expected to show stronger rime effects than German readers. This is exactly what was found, for both word and nonword reading. This preference must have been developmentally established, as similar patterns are found in children (Goswami et al., 1998).

Interestingly, recent simulation work (Perry & Ziegler, 2002) showed that a connectionist learning model (i.e. the dual process model by Zorzi et al. (Zorzi, Houghton & Butterworth, 1998)) could not predict the behavioral patterns found. This was because the greater consistency of the German orthography actually drove the German model to process larger units rather than smaller units. The model adopted the opposite processing pattern to the one observed for the German adults. At the same time, the dual route cascaded model (DRC; Coltheart, Rastle, Perry, Langdon & Ziegler, 2001), which can predict the length effect found in German (i.e. the small grain size effect), could not predict the greater reliance of English adults on orthographic rime units. This is probably because neither phonological nor orthographic processing is sensitive to rime size units in this model. Thus, data such as these seem to suggest that skilled reading cannot be fully understood unless one takes into account the footprints that development leaves on reading.

Implications for theories of reading

Until recently, the most prominent cross-language reading theory has been the orthographic depth hypothesis (ODH; Frost, Katz & Bentin, 1987; Katz & Frost, 1992), which was based on the dual route model of reading (Coltheart et al., 2001). The ODH does not postulate that different psycholinguistic units develop in response...
to differences in orthography. Rather, the ODH suggests that readers adapt their reliance on the ‘orthographic’ (whole word recognition) or ‘phonological’ (recoding) route, depending on the demands of the orthography. In a consistent orthography, readers rely more on the ‘phonological’ or nonlexical route, because the mapping between letters and sounds is relatively direct and unambiguous. In an inconsistent orthography, readers rely less on the phonological route and to a greater extent on the lexical or ‘orthographic’ route.

One key prediction of the ODH is that phonological effects should be reduced in a relatively inconsistent orthography such as English. However, a large number of studies found strong phonological effects in English in a variety of paradigms (Perfetti & Bell, 1991; Rayner, Sereno, Lesch & Pollatsek, 1995; Van Orden, 1987; Ziegler, Van Orden & Jacobs, 1997). Although Katz and Frost correctly point out that such data only challenge the strong version of the ODH, according to which people who read deep orthographies never use phonological information, such studies still show that phonological processes play a role in reading both consistent and inconsistent orthographies. Indeed, some of the original advocates of the ODH have come to the following conclusion: ‘We no longer believe that the difference (between shallow and deep orthographies) is one of whether or not phonology is routinely involved in visual word recognition. . . we now think that the difference is merely methodological, a matter of the greater simplicity with which one can contrive an experimental demonstration of phonological involvement’ (Lukatela & Turvey, 1999, p. 1069).

Given that English seems to lie at the extreme end of the consistency continuum with regard to orthography–phonology relationships, it might even be the case that the prominent dual route architecture (i.e. two separate routes to pronunciation in the skilled reading system) may in fact only develop for English. This is a particularly striking idea given the large number of studies of reading conducted in English, and the influence that theoretical models of reading in English have had on models in other languages.

Importance of grain size, reading strategy and teaching method

How important are considerations of grain size for understanding reading development in different languages? One test is to see whether existing connectionist learning models derived for English can simulate reading development in consistent and inconsistent orthographies without taking differences in grain size, reading strategies or teaching methods into account. To explore this question, Hutzler et al. (Hutzler, Ziegler, Perry, Wimmer & Zorzi, 2004) compared the performance of two major connectionist reading models in two languages, the triangle model (Plaut, McClelland, Seidenberg & Patterson, 1996) and the two-layer associative model (Zorzi et al., 1998). These models were trained on a comparable database of German and English words, and were tested on an identical set of German and English nonwords at different stages during the process of learning to read. The authors found that both models showed an overall advantage for the more consistent orthography (i.e. an advantage for German over English). However, the networks exhibited no cross-language differences during initial learning phases. Rather, there were increasingly large differences during later learning phases. This is the opposite of the empirical pattern (Frith et al., 1998; Goswami et al., 2001), where German beginning readers outperform English beginning readers but differences are attenuating by a reading age of around 10 years.

It seems that the models fail to capture the cross-language learning rate effect because they only deal with the implicit aspects of the learning process. Both models are presented with words that are fully segmented into letters, and they learn about their correspondences with a phonology that is already fully specified in terms of phonemes. In essence, the connection between the two domains is the only thing that is learned. The models behave as if they already contained fully specified orthographic and phonological representations prior to reading. Also, the learning process itself is modelled as beyond the control of the reader or the teacher – it is implicit. In real life, however, learning to read starts out with explicit processes, such as the explicit teaching of small grain-size correspondences. It is these explicit processes and their potential interactions with the more implicit aspects of lexical processing that are missing from the models.

As we have argued, a key feature of learning to read consistent orthographies is the reliability of correspondences at small grain sizes. This boosts the acquisition of phonological recoding and phonemic awareness, especially during the early phases of reading acquisition, and seems to have long-lasting effects on the skilled reading system (Ziegler et al., 2001). Given that current connectionist learning models are not sufficiently sensitive to the fact that literacy acquisition in consistent orthographies starts out with explicit teaching of small-unit correspondences, the failure of these models to fully capture the empirical data is not surprising. In fact, when Hutzler et al. (2004) pre-trained Zorzi’s two-layer associative model on simple grapheme–phoneme correspondences prior to the word learning process, thus
imitating what happens during phonics teaching, the model accurately predicted the cross-language learning rate effect. Together, these data suggest that new connectionist models need to be developed that can encompass critical developmental processes.

Conclusions and future directions

The present review clearly suggests that future researchers will need to integrate domains that have traditionally worked in isolation. Classically researchers have designed their experiments as though skilled reading was unaffected by reading and language development. Here, we suggest that future research needs to construct critical manipulations that can track the mutual dependencies across these domains at different points in development and across different language environments. This is particularly important if the teaching of reading in different languages is to be informed by an effective evidence base. Systematic cross-language research conducted within a psycholinguistic grain size framework is bound to yield rich pedagogical rewards.

References


