Available online at www.sciencedirect.com
SCIENCE DIRECT®
J. Experimental Child Psychology xxx (2005) xxx–xxx
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

The effects of spelling consistency on phonological awareness: A comparison of English and German

Usha Goswami [a,*], Johannes C. Ziegler [b], Ulla Richardson [a,c]

[a] *Faculty of Education, University of Cambridge, Cambridge CB2 2PQ, UK*
[b] *CNRS and Université de Provence, 13331 Marseille, Cedex 3, France*
[c] *Department of Languages/Finnish, University of Jyvaskyla, FIN-40014 Jyvaskyla, Finland*

Received 25 December 2004; revised 26 May 2005

Abstract

Within alphabetic languages, spelling-to-sound consistency can differ dramatically. For example, English and German are very similar in their phonological and orthographic structure but not in their consistency. In English the letter *a* is pronounced differently in the words *bank*, *ball*, and *park*, whereas in German the letter *a* always has the same pronunciation (e.g., *Ball*, *Park*, *Bank*). It is often argued that reading acquisition has a reciprocal effect on phonological awareness. As reading is acquired, therefore, spoken language representation may be affected differently for English and German children. Prior to literacy acquisition, however, phonological representation in English and German children should be similar due to the similar phonological structure of the two languages. We explored this hypothesis by comparing phonological awareness at the rime and phoneme levels in prereaders and beginning readers in English and German. Similar developmental effects were indeed observed in prereaders, but differential effects had emerged within the first year of reading instruction.

* Corresponding author. Fax: +44 1223 324421.
E-mail address: ucg10@cam.ac.uk (U. Goswami).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2005.06.002

Keywords: Phonological representation; Spelling consistency; Rime; Phoneme

Introduction

Reading acquisition in different languages can involve different psycholinguistic grain sizes (e.g., syllables, onsets, rimes, bodies, phonemes) (Ziegler & Goswami, 2005). Within a particular language, phonological awareness at different grain sizes will affect the acquisition of the written language. For example, whereas English and German follow an onset–rime syllable structure, Korean follows a body–coda structure. Onset–rime awareness is linked to reading acquisition in English (e.g., Bradley & Bryant, 1983), whereas body–coda awareness is linked to reading acquisition in Korean (e.g., Yoon, Bolger, Kwon, & Perfetti, 2002). Similarly, the grain sizes captured by different types of spelling units in the written language will affect the representation of the spoken language. Only written languages that represent phonemes via the alphabet seem to promote the development of phoneme awareness in their learners. Whereas children learning to read alphabetic languages with consistent orthographies develop phoneme awareness very rapidly (e.g., Greek: Harris & Giannoulis, 1999; Italian: Cossu, Shankweiler, Liberman, Katz, & Tola, 1988), children learning to read nonalphabetic languages such as Chinese by rote do not (Huang & Hanley, 1994). Because children's awareness of the phonological structure of their spoken language is a crucial determinant of how well they acquire literacy (e.g., Bradley & Bryant, 1983; Bryant, Bradley, & Crossland, 1990; Hoien, Lundberg, Stanovich, & Bjaalid, 1995; Lundberg, Frost, & Petersen, 1988; Schneider, Kuespert, Roth, Vise, & Marx, 1997), it is important to understand cross-language variations in the reciprocal relation between the acquisition of written language and the further development of phonological awareness (e.g., Perfetti, Beck, Bell, & Hughes, 1987; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). In this article, therefore, we compare phonological development before and during the first year of being taught to read for English and German children.

Prior to beginning to learn to read, children are aware of only relatively large phonological grain sizes across languages. In general, they have represented phonological structure at the level of whole words, syllables, onsets, and rimes, with awareness of each grain size developing sequentially in overlapping phases (e.g., Anthony & Lonigan, 2004; Anthony, Lonigan, Driscoll, Phillips, & Burgess, 2003). As they acquire reading, children learning to read alphabetic scripts become aware of the smaller units of sound represented by alphabetic letters, that is, phonemes. As a direct result of being taught letter sound correspondences, children rapidly augment their phonological representations with phoneme-level information (e.g., Durgunoglu & Oney, 1999; Harris & Giannoulis, 1999). It is now widely accepted that interactive relations between spoken and written language characterize the early acquisition of literacy.

Although the developmental sequence of the emergence of phonological awareness (from larger to smaller grain sizes) appears to be universal across languages, there are important differences in the rate of reading acquisition that depend on the consistency of spelling-to-sound relations and on the granularity (i.e., grain size) of orthographic and phonological representations. As already noted, not all languages respect the onset–rime division of the syllable characteristic of European languages. Hence, phonological grain sizes can differ in important ways across languages. Furthermore, the grain size of orthographic units also varies across languages. The critical factor here is the consistency of spelling-to-sound relations, which can be computed statistically. In some writing systems (e.g., English, Danish), letters and letter clusters can have multiple pronunciations. For English, for example, the letter cluster–*ough* has many different spelling-to-sound correspondences (e.g., *bough*, *cough*, *dough*, *tough*, *through*) (Ziegler, Stone, & Jacobs, 1997). In other orthographies (e.g., Greek, Italian, German, Spanish), letters and letter clusters are usually pronounced in the same way. Similarly, in some orthographies, a phoneme can have multiple spellings (e.g., English, French, Hebrew: for English, consider the phoneme /i/ in *leek*, *beak*, *eke*, and *clique*). In other orthographies (e.g., Italian), a phoneme is almost always spelled in the same way. When letter sound relations are very consistent, as in Italian, Greek, and German, it is natural to acquire reading by recoding words at smaller grain sizes. Small grain sizes will be very efficient in such cases because they will almost always yield an accurate recoding of the sounds of the target words. If letters or letter clusters have multiple pronunciations, however, it is natural to look for higher levels of orthographic consistency to supplement phonological recoding at smaller grain sizes. For languages such as English, it may be useful to use rhyme analogies to supplement grapheme–phoneme recoding (Goswami, 1986), particularly for monosyllables, where rimes are more consistent spelling units than single letters (e.g., Treiman, Mullennix, Bijeljac-Babic, & Richmond-Welty, 1995).

Cross-language studies indeed show that grapheme–phoneme recoding skills are acquired rapidly by children learning to read languages such as Greek, Italian, Spanish, and German (e.g., Seymour, Aro, & Erskine, 2003). The accurate decoding of familiar words in these languages approaches ceiling levels within the first year of schooling. English, Danish, and French children progress more slowly, with English children decoding less than 50% of simple familiar words successfully during the first year at school. English and French children also make more use of orthographic rhyme analogies than do Spanish children based on rime and super-rime orthographic units (Goswami, Gombert, & de Barrera, 1998). English children rely on larger spelling units more than do German children, who rely almost exclusively on small grain sizes (Goswami, Ziegler, Dalton, & Schneider, 2001, 2003). Hence, differences in spelling-to-sound consistency result in the use of different orthographic grain sizes by beginning readers in different languages.

It is logically possible that these different grain sizes may have different reciprocal effects on the further development of phonological awareness as reading is acquired. Spelling consistency certainly appears to affect phonological processing in adults. For adults, performance on phonological tasks is more accurate for words that have consistent spelling patterns than for words that have inconsistent spelling patterns across languages (e.g., Ventura, Morais, Pattamadilok, & Kolinsky, 2004; Ziegler & Ferrand, 1998; Ziegler, Ferrand, & Montant, 2004; Ziegler, Muneaux, & Grainger, 2003). For example, adults are faster and more accurate in auditory lexical decision tasks when spoken words contain rimes that are always spelled in the same way throughout the phonological neighborhood (e.g., /uk/ as in *duck, luck*, and *suck*). General orthographic effects in phonological processing tasks have also been demonstrated for adults (Halle, Chereau, & Segui, 2000; Slowiaczek, Soltano, Wieting, & Bishop, 2003; Ziegler & Ferrand, 1998; Ziegler et al., 2004) and for children (e.g., Ehri & Wilce, 1980; Tunmer & Nesdale, 1985). For example, children who can read make orthographically influenced errors when counting phonemes in words. They judge a word such as *pitch* to have more phonemes than a word with a phonologically equivalent rime such as *rich*, or they count twice for the vowel in *look.* Clearly, learning to read has documentable general effects on phonological performance. What remains to be established is whether the degree of consistency in a particular language has specific effects on phonological awareness for words of different types in that language.

In the current study, we investigated both systemwide effects of spelling-to-sound consistency and word-specific effects in English and German. At the systemwide level, the marked spelling-to-sound inconsistency of English was expected to lead to slower development of phoneme awareness by English children. German children encounter largely consistent spelling-to-sound relations in their early reading and so were expected to develop phoneme awareness relatively rapidly. At the word-specific level, we compared the effects of making phonological decisions about words that had either consistent or inconsistent spelling patterns in both languages. A German–English comparison is ideal for this purpose because systemwide spelling-to-sound consistency varies dramatically between the languages, yet the two languages make it possible to come up with equally strong spelling consistency manipulations at the individual word level. Also, it is simple to control for a variety of other factors that are not related to consistency because English and German have a similar orthography and phonology (the languages stem from the same Germanic root).

In this study, therefore, we investigated the effects of learning to read on children's phonological representations at the rime and phoneme levels. We predicted that, prior to learning to read, phonological development in the two languages should be very similar. This is because English and German share a highly similar phonological structure and because rime neighbors dominate within phonological neighborhoods in both languages despite the greater absolute number of rime neighbors (roughly four times more) in English (Ziegler & Goswami, 2005). However, once children had been learning to read for roughly a year, a divergence in phonological performance was expected by language. If spelling-to-sound consistency affects phonological development in a systemwide manner once reading has begun, then German children might show enhanced phoneme awareness. Words that are consistently spelled should be processed more accurately, and more words encountered by German children will have consistent spellings. At the word-specific level, however, particular neighborhoods that have inconsistent members in both English and German can be selected. For example, to judge that *bank* and *tank* rhyme (consistent neighborhood: German = *Bank* and *Tank*) might be easier than judging that *boat* and *note* rhyme (inconsistent neighborhood: German = *Boot* and *Not*) for children learning to read both languages.

To compare phonological development in English and German children, we used two versions of the oddity task pioneered by Bradley and Bryant (1978). One was rime oddity (Experiment 1), and the other was vowel oddity (Experiment 2). In the oddity task, the children listened to triplets of words. For each oddity trial, the children had to select the odd word out. In Experiment 1 the odd word had a different rime (e.g., *house*, *mouse*, ***kiss***), and in Experiment 2 the odd word had a different vowel phoneme (e.g., *house*, *loud*, ***path***). The spelling consistency of the word triplets was manipulated in each experiment. Half of the words were consistent in that their phonological rime (Experiment 1) or vowel (Experiment 2) was spelled in the same way throughout the relevant neighborhood. Half of the words were inconsistent in that their phonological rime (Experiment 1) or vowel (Experiment 2) could be spelled in multiple ways (e.g., *word* vs. *bird*). Among these inconsistent triplets, half of the items selected did not share the same spelling for the body (*word/bird:* inconsistent, different rime spelling in exemplars chosen), and the other half did share the same spelling for the body (*good/hood:* inconsistent, same rime spelling in exemplars chosen). In the first case, if orthographic information is activated in an on-line fashion, then the–*ord* spelling of *word* will conflict with the–*ird* spelling of *bird*, producing what might be called an on-line consistency effect. In contrast, in the case of inconsistent triplets with identical spellings (*good/hood*), this on-line conflict should be absent. Thus, finding a consistency effect in the latter condition would suggest that inconsistent words have less stable or accurate phonological representations per se. Such an effect could be interpreted as a residual consistency effect (Muneaux & Ziegler, 2004).

Experiment 1

Method

Participants

Details of the participants are given in Table 1. Participants were recruited in Great Britain and Germany. Although race and social class data were not collected systematically, the majority of participants were of Caucasian descent and were from middle-class neighborhoods. In both countries, a group of prereaders was selected in kindergarten. These children were matched in age and had not yet received any formal reading instruction. To be included in the study as prereaders, English children had to read fewer than 10 words on the British Ability Scales (BAS) (Elliott, Smith, & McCulloch, 1996). Seven children were not able to read any words, and the median number of words read was 3. The same criterion was applied to the German children; they had to read fewer than 10 words on the Würzburger Leise Leseprobe (WLLP) (Küspert & Schneider, 1998). Letter names and letter sounds are not taught in German kindergarten; consequently, letter name and letter sound knowledge was much better in the English prereaders. German children enter school 1 year later than do English children (i.e., roughly 6 years 6 months of age). Initially, we started testing German children in first grade. However, at the time when we tested (3 months after

Table 1
Characteristics of participating children in Germany and Great Britain

	Prereaders			Readers	
	English	German 5-year-olds	German 6-year-olds	English	German
Sample size	17 (9 girls and 8 boys)	12 (7 girls and 5 boys)	12 (7 girls and 5 boys)	18 (11 girls and 7 boys)	12 (5 girls and 7 boys)
Age (range)	5 years 2 months (3 years 8 months to 5 years 9 months)	5 years 2 months (4 years 3 months to 5 years 8 months)	6 years 9 months (6 years 5 months to 7 years 6 months)	6 years 2 months (5 years 3 months to 6 years 7 months)	7 years 7 months (7 years 0 months to 8 years 5 months)
Reading level	< 10[a]	< 10[b]	< 10	111 (89–140)[c]	54 (14–91)[d]
Letter name	22.6	2.9	26.0	25.9	26.0
Letter sound	18.2	0.7	25.1	25.8	26.0

[a] Number of words read on the British Ability Scales.
[b] Number of words read on the Würzburger Leise Leseprobe.
[c] Standard reading scores on the British Ability Scales.
[d] Percentile reading scores on the Würzburger Leise Leseprobe.

the beginning of the school year), German first graders had not yet started to read words. They were still learning letter names and sounds and some basic grapheme–phoneme correspondences. Indeed, all of the first graders still read fewer than 10 words of the WLLP. Thus, it was not possible to match the German and English participants for both reading level and age. Instead, we selected an additional group of older German children in early second grade who were matched to the English readers in terms of reading ability. Both groups of readers were slightly ahead of their age norm (English norm = 100, German norm = 50).

Stimuli

A total of 48 word triplets, 24 in each language, were selected. As in our previous work (Ziegler, Perry, Jacobs, & Braun, 2001), we tried to select cognates whenever this was possible (e.g., *storm* in English vs. *Sturm* in German). As required for the rime oddity task, two of the three words shared the same rime and the third word did not (i.e., the odd word out). Of these phonologically similar pairs, half were inconsistent (i.e., their rimes could be spelled in multiple ways) and the other half were consistent (i.e., their rimes could be spelled in only one way). In half of the inconsistent triplets, the inconsistent words shared the same rime spelling (residual consistency effect); in the other half, the inconsistent words had different rime spellings (on-line consistency effect). All items are listed in Appendix A.

Consistency was calculated using the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). The same procedure was used as in previous consistency analyses (Ziegler, Jacobs, & Stone, 1996; Ziegler et al., 1997). That is, for each word, we calculated its consistency ratio, which reflects the number of friends divided by the number of friends and enemies. For this analysis, friends were defined as words in which the phonological rime was spelled the same, whereas enemies were defined as words in which the rime was spelled differently. Note that we were able to come up with equally strong consistency manipulations across the different languages. This can be seen in the comparable size of the consistency ratios. Examples and item statistics are given in Table 2.

Because phonological neighborhood has been shown to influence phonological development (Metsala, 1997) and performance in a variety of phonological tasks (De Cara & Goswami, 2002, 2003; Goldinger, Luce, & Pisoni, 1989; Luce, Pisoni, & Goldinger, 1990; Ziegler et al., 2003), we also matched neighborhood density across languages (Table 2). Finally, although we used cognates to control for potential frequency differences across languages, we looked up the CELEX frequencies of words in the two languages. We obtained very similar frequency estimates across languages and conditions (all Fs < 1).

Task administration

The words for each oddity trial were recorded by a native female speaker of British English and a native male speaker of German and then were digitized for computer presentation using Cool Edit 96 (Syntrillium Software). Before each trial, the

Table 2
Item characteristics (means and standard deviations) of stimuli in Experiment 1

	English			German		
	Consistent	Inconsistent/ Same spelling	Inconsistent/ Different spelling	Consistent	Inconsistent/ Same spelling	Inconsistent/ Different spelling
Number of words	24	12	12	24	12	12
CELEX frequency[a]	248 (313)	276 (267)	199 (295)	397 (242)	358 (191)	356 (326)
Rime neighbors	12.5 (6.7)	14.5 (9.2)	14.5 (7.9)	10.1 (4.3)	10.1 (4.3)	8.8 (6.1)
Consistency ratio	.96 (.05)	.41 (.12)	.46 (.05)	.97 (.05)	.50 (.17)	.46 (.05)
Example	house/mouse	half/calf	boat/note	Haus/Maus	Halb/Kalb	Boot/Not

Note. Standard deviations are in parentheses.
[a] Frequency per million. Frequencies greater than 1000 were truncated.

children saw in the center of the computer screen a row of asterisks that disappeared when the trial began. The screen then went blank. The stimuli were presented through headphones. The children had to say the word when they knew the answer. Trials were not blocked by consistency; rather, three different semirandomized sequences of the word triplets, which also varied the position of the odd word in each triplet, were created. Each participant received one of the possible sequences, and the three sequences were counterbalanced across participants.

Results

Performance in the rime version of the oddity task is presented in Table 3. The overall consistency effect can be quantified by contrasting inconsistent words (aver-

Table 3
Mean performances (percentages correct) in the rime version of the oddity task for prereaders and readers in Great Britain and Germany

	Condition	Prereaders			Readers	
		English	German 5-year-olds	German 6-year-olds	English	German
Consistent	1	62.3	49.7	68.1	89.8	86.8
Inconsistent overall	2	59.8	56.9	63.2	77.8	85.4
Inconsistent/Different spellings	3	55.9	59.7	62.5	70.4	86.1
Inconsistent/Same spellings	4	63.7	55.3	63.9	85.2	84.7
Overall consistency effect	1–2	2.5	−7.2	4.9	12.0*	1.4
On-line consistency effect	1–3	6.4	−10.0	5.6	19.5**	0.7
Residual consistency effect	1–4	−1.4	−5.6	4.2	4.6	2.1

* $p < .001$.
** $p < .0001$.

aging across same and different spellings) with consistent words. The on-line consistency effect can be estimated by contrasting inconsistent words with different spellings for the exemplars chosen from the neighborhood against consistent words. Finally, the residual consistency effect can be estimated by contrasting inconsistent words with identical spellings for the exemplars chosen from the neighborhood against consistent words (for subtractions, see Table 3). We conducted three analyses of variance (ANOVAs) to assess the different types of consistency effect—overall, on-line, and residual—and their potential interactions with language (German vs. English) and group (reader vs. prereader). ANOVAs were performed with participants (F_1) and items (F_2) as the random factor. In the analysis by participants, consistency was a within-subject factor, whereas language and group were between-subject factors. In the analyses by items, consistency and language were between-item factors, whereas group was a within-item factor because readers and prereaders were presented with the same items.

Overall consistency

The three-factorial ANOVAs (Consistency × Group × Language) showed a significant main effect of group, $F_1(1, 55) = 27.2$, $p < .0001$, $F_2(1, 44) = 169.3$, $p < .0001$, and no significant main effect of language, $F_1(1, 55) = 0.26$, $F_2(1, 44) = 1.8$, or overall consistency, $F_1(1, 55) = 1.1$, $F_2(1, 44) = 2.3$. Significant two-way interactions were obtained between consistency and language, $F_1(1, 55) = 6.1$, $p < .05$, $F_2(1, 44) = 4.6$, $p < .05$, and between consistency and group, $F_1(1, 55) = 4.9$, $p < .05$, $F_2(1, 44) = 7.6$, $p < .01$. No other interactions reached significance, all Fs < 1.

The main effect of group indicates that readers obtained higher scores than did prereaders. The absence of a language main effect suggests that, overall, English and German children obtained comparable scores in the rime task. The interaction between the effects of consistency and language reflects the fact that a consistency effect was observed for the English children but not for the German children. The interaction between consistency and group reflects the fact that consistency effects were observed for readers but not for prereaders. Post hoc comparisons using two-tailed Student t tests confirmed that only English readers showed a significant consistency effect (Table 3).

On-line consistency

The statistical pattern was very similar to that found in the overall consistency analysis. We found a significant main effect of group, $F_1(1, 55) = 24.2$, $p < .0001$, $F_2(1, 32) = 81.5$, $p < .0001$, and no significant main effect of language, Fs < 1. The on-line consistency effect was significant by items but not by participants, $F_1(1, 55) = 2.6$, $p < .11$, $F_2(1, 32) = 5.1$, $p < .05$. As in the previous analysis, the main effects were qualified by significant interactions between on-line consistency and language, $F_1(1, 55) = 11.9$, $p < .001$, $F_2(1, 32) = 15.6$, $p < .001$, and between on-line consistency and group, $F_1(1, 55) = 5.4$, $p < .05$, $F_2(1, 32) = 3.7$, $p < .06$. The triple interaction failed to reach significance, Fs < 1. Again, consistency effects were observed for readers but not for prereaders.

Residual consistency

Apart from a significant main effect of group, $F_1(1, 55) = 25.1$, $p < .0001$, $F_2(1, 32) = 118.1$, $p < .0001$, and a significant main effect of language by items, $F_2(1, 32) = 8.4$, $p < .01$ (English children were more accurate than German children in judging rimes), no other main effects or interactions reached significance, all $F\text{s} < 1.6$.

German 6-year-olds versus English 6-year-olds

The previous analyses suggest a consistency effect in the rime task for English readers only that has its basis in on-line processing. However, the English readers differed in age from the German readers. Therefore, we ran a separate comparison between the German first- grade prereaders and the English first-grade readers. In this comparison, all participants were 6 years of age. If the emergence of a consistency effect can be explained by maturational changes in the language system, then the two groups should show similar performance. A 2×2 ANOVA with language and consistency as factors showed main effects of language (by subjects only, English children were more accurate than German children), $F_1(1, 28) = 6.7$, $p < .05$, $F_2(1, 44) = 1.2$, and consistency, $F_1(1, 28) = 11.9$, $p < .01$, $F_2(1, 44) = 6.2$, $p < .01$. There was also a significant interaction by items between consistency and language, $F_1(1, 28) = 2.1$, $p < .15$, $F_2(1, 44) = 4.0$, $p < .05$. The interaction arose because the English children showed a significant consistency effect (90% accuracy for consistent items, 78% accuracy for inconsistent items), whereas the German prereaders of the same age did not (68% accuracy for consistent items, 64% accuracy for inconsistent items).

The residual versus on-line consistency analyses confirmed that the source of this consistency effect for the English children was on-line processing. The on-line consistency ANOVA (i.e., comparing inconsistent words with different spellings against consistent words) showed significant main effects of language, $F_1(1, 28) = 3.7$, $p < .06$, $F_2(1, 36) = 5.4$, $p < .05$, and consistency, $F_1(1, 28) = 11.9$, $p < .01$, $F_2(1, 36) = 10.7$, $p < .01$, and a significant interaction between language and consistency, $F_1(1, 28) = 3.7$, $p < .06$, $F_2(1, 36) = 9.4$, $p < .01$. The English children made accurate rhyme judgments about 90% of the consistent items compared with 70% of the inconsistent items that were spelled differently Comparable figures for the residual analysis were 90 and 85% (a nonsignificant difference). In contrast, the German children made accurate rhyme judgments about 68% of the consistent items compared with 63% of the inconsistent items that were spelled differently Comparable figures for the residual analysis were 68 and 64% (both nonsignificant). In the residual consistency analyses, the critical consistency by language interaction failed to reach significance, all $F\text{s} < 1$.

Discussion

The results of Experiment 1 demonstrate spelling consistency effects in phonological processing tasks in children. Comparable experiments manipulating rime-level consistency have previously demonstrated such consistency effects in adults

(Halle et al., 2000; Seidenberg & Tanenhaus, 1979; Ventura, Kolinsky, Brito-Mendes, & Morais, 2001; Ziegler & Ferrand, 1998; Ziegler et al., 2004). That the consistency effect is present in child readers but not in child prereaders provides a striking demonstration that the effect is a direct consequence of learning to read. The absence of the consistency effect in prereaders of different ages rules out explanations based on spoken language development. Note also that the similarity in absolute levels of performance shown by the prereaders in the two languages occurred despite quite different levels of letter name and letter sound knowledge. This questions the idea that early phonological awareness cannot be distinguished from early alphabet knowledge (Castles & Coltheart, 2004).

The consistency effect for the English readers was present in the overall consistency comparison, which added data across individual items from the same inconsistent neighborhood irrespective of the spellings of the particular items chosen. However, further analyses comparing on-line effects of consistency with structural–residual ones demonstrated repeatedly that the effect was significant only when consistent words were compared with inconsistent words with different rime spellings (e.g., *bird/ word*). This suggests that participants were activating the spellings of the words they heard in an on-line fashion and that their phonological decisions were based at least partially on this on-line activation. According to this automatic activation view, two inconsistent words that share different rime spellings will suffer from on-line competition between competing spellings even if a purely phonological judgment is required.[1] When two words from an inconsistent neighborhood share the same spelling, on-line competition is reduced because the spelling enemy is not explicitly present. However, residual effects may become stronger in older children and adults. Consistency effects may be observed in conditions where spelling enemies are not explicitly present when older participants are studied (Muneaux & Ziegler, 2004). This could reflect the residual effects of consistency on building high-quality phonological representations.

It is theoretically interesting, although unexpected, that the consistency effect was present only for the English readers. One possible explanation is that differences in teaching methods for initial reading in the two countries lie behind this difference; that is, teaching methods used with English children may focus more on spelling patterns for rimes. However, this seems unlikely given that the English children in this study were following the National Literacy Strategy (NLS). During the first year of reading instruction, the NLS places emphasis on decoding via applying grapheme–phoneme correspondences. Direct instruction about rimes does not come until the second year of the NLS. Hence, the types of instruction received during the first year of direct teaching of reading in the two countries are similar, focusing on grapheme–phoneme correspondence. An alternative explanation for the difference by language is that only rime consistency was manipulated in Experiment 1. The word triplets with consistent rime spellings had only a single possible correspondence at this level,

[1] It should be noted, however, that we did not test the children's spelling knowledge of each word used in the oddity tasks.

as was the case in previous adult studies (Stone, Vanhoy, & Van Orden, 1997; Ziegler & Ferrand, 1998; Ziegler, Montant, & Jacobs, 1997). However, it could be the case that Germans are less sensitive to rimes as psycholinguistic units, either in phonology or in learning to read, because the absolute size of the monosyllabic lexicon is so much smaller in German. Comparable claims have been made, for instance, for French (Peereman, Content, & Bonin, 1998). In this case, Germans may show consistency effects for smaller grain sizes, namely single phonemes (Perry, 2003). To address this issue, we designed Experiment 2. Here we used a phoneme version of the oddity task.

Experiment 2

For Experiment 2, we created a vowel version of the oddity task by reassembling most of the word triplets used in the rime oddity task. This meant that whereas the previous experiment manipulated phonological similarity at the rime level (e.g., *book*, *took*, *dark*), the current experiment manipulated phonological similarity at the phoneme level (e.g., *book*, *hood*, *ten*). The reassembly strategy made it unlikely that potential performance differences across experiments could be due to item familiarity. Instead, any differences found should be due to the linguistic level of the phonological manipulation.

Method

Participants

The participants were the same as in Experiment 1.

Stimuli and task administration

We constructed 40 new triplets, 20 in each language, using most of the stimuli of Experiment 1. All items are listed in Appendix B. In each triplet, two words shared the same vowel. One word in the triplet shared no phonemes with the other two words and was the odd one out. Half of the words that shared the vowel had a vowel that could be spelled in multiple ways (inconsistent triplets), whereas the other half had a vowel that could be spelled in only one way (consistent triplets). Of the inconsistent triplets, half of the exemplars chosen shared the same vowel spelling (residual consistency effect), whereas the other half had a different vowel spelling (on-line consistency effect). As before, consistent and inconsistent pairs were matched in terms of word frequency and neighborhood size, $F < 1$. Item characteristics can be found in Table 4. Tasks were administered as in Experiment 1.

Results and discussion

Children's data are presented in Table 5. As in Experiment 1, the data were submitted to three pairs of $2 \times 2 \times 2$ ANOVAs testing for the effects of overall, on-line,

Table 4
Item characteristics (means and standard deviations) of stimuli in Experiment 2

	English			German		
	Consistent	Inconsistent/ Same spelling	Inconsistent/ Different spelling	Consistent	Inconsistent/ Same spelling	Inconsistent/ Different spelling
Number of words	20	10	10	20	10	10
CELEX frequency[a]	242 (454)	193 (146)	274 (233)	371 (316)	439 (269)	371 (213)
Rime neighbors	11.5 (5.7)	12.5 (7.2)	13.1 (7.9)	11.1 (5.3)	12.1 (5.3)	10.8 (6.1)
Consistency ratio	.86 (.09)	.32 (.18)	.13 (.12)	.97 (.05)	.72 (.06)	.12 (.09)
Example	half/barn	green/sweet	dawn/fork	Halb/Barsch	Plan/Graf	Clown/Haus

Note. Standard deviations are in parentheses.
[a] Frequency per million, Frequency greater than 1000 truncated.

Table 5
Average performances (percentages correct) in the vowel version of the oddity task for prereaders and readers in Great Britain and Germany

	Condition	Prereaders			Readers	
		English	German 5-year-olds	German 6-year-olds	English	German
Consistent	1	38.8	47.5	48.3	56.7	92.5
Inconsistent overall	2	39.4	52.5	48.8	55.0	81.7
Inconsistent/Different spellings	3	41.2	50.0	47.2	54.4	78.3
Inconsistent/Same spellings	4	37.6	55.6	50.0	55.6	85.0
Overall consistency effect	1–2	−0.6	−5.0	−0.5	1.7	10.8*
On-line consistency effect	1–3	−2.4	−2.5	1.1	2.3	14.2*
Residual consistency effect	1–4	1.2	−8.1	−1.7	1.1	7.5

Note. Stimuli were either consistent, inconsistent with identical vowel spelling, or inconsistent with different vowel spellings.
* $p < .05$.

and residual consistency across languages (German vs. English) and reader groups (prereader vs. reader) by subjects and by items. The ANOVAs were identical in design to those in the previous experiment.

Overall consistency

The three-factorial ANOVAs (Consistency × Group × Language) showed significant main effects of group, $F_1(1, 55) = 37.2$, $p < .0001$, $F_2(1, 36) = 131.3$, $p < .0001$, and language, $F_1(1, 55) = 22.8$, $p < .0001$, $F_2(1, 36) = 48.1$, $p < .0001$. No significant main effect was obtained for consistency, both Fs < 1. Significant effects were obtained for the two-way interactions between group and language, $F_1(1, 55) = 5.3$, $p < .05$, $F_2(1, 36) = 21.4$, $p < .0001$, and consistency and group, $F_1(1, 55) = 2.8$, $p < .10$. $F_2(1, 36) = 4.2$, $p < .05$. The triple interaction failed to reach significance, $F_1(1, 55) = 1.6$, $F_2(1, 36) = 2.5$, $p < .12$.

The main effect of group confirms that readers performed better than prereaders in the vowel task. The main effect of language arose because the German children performed better than the English children. These main effects were qualified by a

Group × Language interaction that arose because the German readers outperformed the English readers, but the prereaders in the two languages were similar. The Consistency × Group interaction indicated the presence of an overall consistency effect that was restricted to the readers. Indeed, post hoc analyses confirmed that only the German readers showed significant consistency effects in the vowel task (Table 5).

On-line consistency

The ANOVAs revealed a very similar pattern to the overall consistency analysis. We found significant main effects for group, $F_1(1, 55) = 27.7$, $p < .0001$, $F_2(1, 26) = 90.6$, $p < .0001$, and language, $F_1(1, 55) = 15.1$, $p < .0001$, $F_2(1, 26) = 25.9$, $p < .0001$, indicating that readers were better than prereaders and that German children performed better than English children. The main effect of consistency failed to reach significance, both Fs < 1. There was a significant interaction between group and language, $F_1(1, 55) = 4.5$, $p < .05$, $F_2(1, 26) = 17.2$, $p < .0001$. This arose because the German readers, but not the prereaders, outperformed the English groups. The Consistency × Group interaction just failed to reach significance, $F_1(1, 55) = 2.6$, $p < .11$. $F_2(1, 26) = 3.5$, $p < .07$. However, the critical comparison between vowel judgments for the German readers was significant in post hoc tests (93% correct for consistent words vs. 82% correct for inconsistent words, $p < .05$). No other interaction reached significance, all Fs < 1.

Residual consistency

The ANOVAs exhibited significant main effects of group, $F_1(1, 55) = 34.9$, $p < .0001$, $F_2(1, 26) = 89.1$, $p < .0001$, and language, $F_1(1, 55) = 24.3$, $p < .0001$, $F_2(1, 26) = 36.0$, $p < .0001$, and a significant interaction between language and group, $F_1(1, 55) = 4.3$, $p < .05$. $F_2(1, 26) = 12.3$, $p < .01$. None of the remaining effects reached significance.

German 6-year-olds versus English 6-year-olds

As in Experiment 1, we performed a 2 × 2 ANOVA with language and consistency as factors using only the data of the German and English 6-year-olds. The ANOVA showed that neither group nor language had a significant effect, all Fs < 1. Also, there was no interaction between these factors, $F < 1$.

Cross-task comparisons

For vowel judgments, therefore, consistency effects are limited to German readers and appear to arise from on-line processing. For rime judgments (Experiment 1), consistency effects are limited to English readers and again appear to arise from on-line processing. However, to explore the potential systemwide effects of spelling-to-sound consistency, we need to compare language-specific patterns of performance in the rime and phoneme oddity tasks as the children learn to read. This makes it inter-

Table 6
Cross-task comparison between performance in the rime task of Experiment 1 and performance in the vowel task of Experiment 2 (percentages)

	Prereaders			Readers	
	English	German 5-year-olds	German 6-year-olds	English	German
Rime task	61	53	66	84	86
Vowel task	39	50	48	56	87
Task difference	22**	3	18*	28**	−1

* $p < .001$.
** $p < .0001$.

esting to run a cross-task comparison combining the data from the two experiments. Group performance is shown in Table 6. These data were submitted to 2 × 2 × 2 ANOVAs using task (rime vs. vowel), group (reader vs. prereader), and language (English vs. German) as factors. In the analysis by participants, task was a within-subject variable, whereas group and language were between-subject variables. In the analysis by items, group was a within-item variable because both readers and prereaders were presented with the same items, whereas language and task were between-item variables.

The ANOVAs exhibited significant effects of task, $F_1(1, 55) = 28.6$, $p < .0001$, $F_2(1, 84) = 59.8$, $p < .0001$, group, $F_1(1, 55) = 42.3$, $p < .0001$, $F_2(1, 84) = 24.9$, $p < .0001$, and language, $F_1(1, 55) = 4.7$, $p < .05$, $F_2(1, 84) = 24.9$, $p < .0001$. These main effects reflect the fact that the rime task was easier than the vowel task, that readers performed better than prereaders, and that German children obtained higher scores than English children. The main effects were qualified by two significant interactions. The interaction between language and group was significant, $F_1(1, 55) = 3.3$, $p < .08$, $F_2(1, 84) = 25.9$, $p < .0001$, reflecting the fact that the largest cross-language task difference was obtained for the readers. The German readers made accurate judgments on 87% of the trials compared with 70% of the trials for the English readers. Importantly, there was also a significant task by language interaction, $F_1(1, 55) = 23.3$, $p < .0001$, $F_2(1, 84) = 45.1$, $p < .0001$. This arose because performance on the vowel task was quite similar to performance on the rime task for the German children (62% accurate judgments vs. 68% accurate judgments, respectively) but not for the English children (48% accurate judgments vs. 73% accurate judgments, respectively). The English children, overall, showed a much larger accuracy difference in making similarity judgments about rimes versus phonemes.

As in the previous experiment, an additional ANOVA was performed to control for age, comparing the German 6-year-olds who were prereaders with the English 6-year-olds who were readers. The results were nearly identical to those of the previous ANOVA except that now the triple interaction among language, group, and task was significant, $F_1(1, 55) = 6.2$, $p < .05$, $F_2(1, 84) = 12.7$, $p < .0001$. The triple interaction reflects the fact that both the German and English children showed better performance on the rime task than on the vowel task (66% accuracy vs. 48% accuracy for German, 84% accuracy vs. 56% accuracy for English). However, this difference was significantly larger for the English children (28% for English vs. 18% for German). As

Table 6 shows, once German children learn to read, rimes rapidly lose their transient advantage as salient phonological units.

General discussion

In the current study, we explored the impact of spelling consistency on the development of phonological awareness at both a systemwide and a word-specific level in English and German. Psycholinguistic grain size theory (Ziegler & Goswami, 2005) was used as a theoretical framework concerning the expected pattern of results. We expected that, prior to reading, German and English children would show very similar phonological development, characterized by good phonological awareness for rimes but not for phonemes. We also expected no effects of spelling-to-sound consistency on phonological awareness prior to literacy. We indeed found that spelling-to-sound consistency did not affect phonological development in prereaders. In fact, the absolute level of phonological awareness was very similar across orthographies. However, the English prereaders showed much stronger rime effects than did the German prereaders. The former displayed a 22% advantage in the rime version of the oddity task compared with a 3% advantage for the Germans.

As the children were taught to read, we expected to find a divergence between languages due to the systemwide differences in spelling-to-sound consistency between German and English. In particular, we expected the German children to become sensitive to phoneme units more quickly than the English children due to the consistency of small reading units (grapheme–phoneme correspondences) in German. English children cannot use a small-unit reading strategy as efficiently as can Germans because spelling-to-sound inconsistency is maximal for small units in English (Treiman et al., 1995). The development of phoneme awareness, therefore, might lag behind that of German children. Consistent with this prediction, we found that after only 1 year of reading instruction, phoneme awareness improved dramatically (by roughly 40%) for the German children only. Indeed, after 1 year of instruction, the German readers were approaching ceiling on both the rime and vowel tasks, whereas the English readers were approaching ceiling on the rime task only.

In terms of word-level effects in the two languages, the data showed that the impact of spelling consistency occurred at different grain sizes for English and German. We had expected that words with inconsistent orthographic patterns would be processed less accurately in a phonological task than would words with consistent orthographic patterns. Our data showed that this was indeed the case, but at the phoneme level for the Germans and at the rime level for the English. The German readers made fewer accurate phonological judgments for words with inconsistent spellings for particular phonemes, whereas the English readers made fewer accurate phonological judgments for words with inconsistent spellings for particular rimes. These interactions between orthographic consistency and phonological grain size suggest that consistency affects only the most salient phonological units within each language. If we had looked for main effects of consistency without taking phonological grain size differences into account, the results would have been seriously misleading. Furthermore, this effect was present only in readers in both languages. This suggests that the source of the consistency effect in phonological processing is indeed orthographic. As discussed previously, the data also suggest that consistency effects in the current study are due to on-line competition from competing spellings rather than to structural–residual effects within the lexicon.

The only difference by grain size in phonological awareness prior to reading was the lack of a rime superiority effect in the German prereaders. This is difficult to explain. Rime neighbors are more frequent than other types of neighbors in both languages. However, neighborhood structure is only one factor that may affect the development of phonological awareness at different grain sizes. Other potentially important factors include the absolute number of neighbors that make a particular psycholinguistic grain size salient and the rate at which words occur in the spoken language. For example, the English monosyllabic lexicon is at least four times as large as the German monosyllabic lexicon, and monosyllabic words make up approximately 90% of all word occurrences in English. This constant exposure to monosyllables may help English children to develop sensitivity to the rime as a psycholinguistic unit. In addition, factors controlled in this experiment, such as age of acquisition, word frequency, and letter sound knowledge, might be expected to affect the development of phonological awareness (Garlock, Walley, & Metsala, 2001; Mann & Wimmer, 2002; Metsala, 1997; Metsala & Walley, 1998). Further research is needed to see whether the absolute size of the set of relevant tokens in a language affects phonological development in the similarity-based way suggested here. This proposal seems plausible given that absolute set size does affect statistical learning in studies of adults (e.g., Homa, Dunbar, & Nohre, 1991), and the number of available exemplars is known to affect similarity-based learning by analogy at the conceptual level (Goswami, 1992).

The current results, although largely supporting psycholinguistic grain size theory (Ziegler & Goswami, 2005), show clearly the need to track the mutual developmental dependencies between written and spoken language skills from a more dynamic and interactive perspective (Van Orden, Jansen op de Haar, & Bosman, 1997). Phonological sensitivity prior to reading differs across languages, and these differences seem to have an important influence on the kinds of units that will play a role during learning to read and write. In turn, learning to read and write affects spoken language development in ways that had not been anticipated by classic theories of auditory word recognition and phonological development. The presence of robust word-specific spelling consistency effects in both languages suggests that the development of phonological representations is not stage-like but rather dynamic and item specific (see also Metsala, 1997; Metsala & Walley, 1998). Further cross-language data using languages with different phonological structures (e.g., Korean, Japanese) are now needed to test the generality of psycholinguistic grain size theory. Such systematic cross-language research is imperative for a deeper understanding of both normal and impaired reading and spoken language development.

Acknowledgment

We thank the head teacher, teachers, and children of Thornhill Primary School (London), Panshanger Primary and Applecroft Primary Schools (Welwyn Garden City, UK), and Dietrich-Bonhoefferfor School (Dietzenbach, Germany) for taking part in this study. We also thank Andy Faulkner and Jill House for their help in preparing the digitized speech stimuli. Our collaboration was supported by an Alliance travel grant (PN 00.230). Support for this research was provided by an ESRC grant (RN 000 239084) to Usha Goswami. Ulla Richardson was a research fellow on the Faculty of Education at Cambridge University when this work was carried out.

Appendix A. Stimuli of Experiment 1

English			German		
consistent triplets					
bank	tank	pink	bank	tank	pink
book	took	dark	buch	tuch	dach
wine	mine	son	wein	mein	sohn
house	mouse	kiss	haus	maus	kuss
round	found	kind	rund	fund	kind
trick	click	stuck	trick	knick	stück
when	then	chin	wann	dann	kinn
brave	grave	drive	geist	meist	fest
brown	clown	swan	ball	fall	heil
bus	plus	cross	horn	dorn	fern
kit	hit	nut	gut	hut	laut
bath	path	mouth	warm	farm	lärm
inconsistent/different rime spellings					
boat	note	root	boot	not	rat
word	bird	hard	wort	bord	hart
state	great	fleet	staat	grad	flut
horn	dawn	fern	braun	clown	schwan
red	dead	wood	brav	graf	drauf
warm	form	lime	grün	kühn	ton
inconsistent/same rime spellings					
half	calf	roof	halb	kalb	gelb
ball	fall	heal	bad	pfad	mut
ghost	most	feast	bus	plus	gross
green	keen	tone	kran	plan	stein
good	hood	loud	kalt	halt	mild
crane	plane	stone	rot	tot	wut

Appendix B. Stimuli of Experiment 2

English			German		
consistent triplets					
when	fell	thick	wann	fall	dick
ball	walk	sight	ball	wand	sich
mine	side	wall	bank	lamm	reif
ghost	rope	bush	geist	seil	busch
trick	stitch	flood	trick	stich	flut
half	barn	foot	halb	brasch	fuss
kit	tip	shell	kalt	hast	schiff
bus	nut	dark	mein	seit	wall
brave	state	plus	farm	salz	pink
plane	grave	sitck	maus	laut	pfad
inconsistent/different vowel spellings					
great	stale	click	boot	lohn	fell
form	salt	pink	kühn	süss	not
word	firm	tooth	grad	stahl	knick
dawn	fork	rice	clown	haus	fisch
clown	house	fish	brav	staat	plus
inconsistent/same vowel spellings					
boat	loan	fall	rot	los	wein
green	sweet	blood	buch	hut	zehn
bath	sharp	full	gut	tuch	laib
book	hood	ten	plan	graf	stock
bank	lamb	ripe	bad	schaf	voll

References

Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology, 96*, 43–55.

Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M., & Burgess, S. R. (2003). Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly, 38*, 470–487.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bradley, L., & Bryant, P. E. (1978). Difficulties in auditory organization as a possible cause of reading backwardness. *Nature, 271*, 746–747.

Bradley, L., & Bryant, P. E. (1983). Categorising sounds and learning to read: A causal connection. *Nature, 310*, 419–421.

Bryant, P. E. M., Bradley, L., & Crossland (1990). Rhyme, alliteration, phoneme detection, and learning to read. *Developmental Psychology, 26*, 429–438.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*, 77–111.

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9*, 1–16.

De Cara, B., & Goswami, U. (2002). Similarity relations among spoken words: The special status of rimes in English. *Behavior Research Methods, Instruments, & Computers, 34*, 416–423.

De Cara, B., & Goswami, U. (2003). Phonological neighborhood density: Effects in a rhyme awareness task in five-year-old children. *Journal of Child Language, 30*, 695–710.

Durgunoglu, A. Y., & Oney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading & Writing, 11*, 281–299.

Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics, 1*, 371–385.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales* (2nd ed.). Windsor, UK: NFER–Nelson.

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language, 45*, 468–492.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language, 28*, 501–518.

Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42*, 73–83.

Goswami, U. (1992). *Analogical reasoning in children*. Hove, UK: Psychology Press.

Goswami, U., Gombert, J. E., & de Barrera, L. F. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics, 19*, 19–52.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *Journal of Memory and Language, 45*, 648–664.

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.

Halle, P. A., Chereau, C., & Segui, J. (2000). Where is the /b/ in "absurde" [apsyrd]? It is in French listeners' minds. *Journal of Memory and Language, 43*, 618–639.

Harris, M., & Giannoulis, V. (1999). Learning to read and spell in Greek: The importance of letter knowledge and morphological awareness. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 51–70). Cambridge, UK: Cambridge University Press.

Hoien, T., Lundberg, L., Stanovich, K. E., & Bjaalid, I. K. (1995). Components of phonological awareness. *Reading & Writing, 7*, 171–188.

Homa, D., Dunbar, S., & Nohre, L. (1991). Instance frequency, categorization, and the modulating effect of experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*, 444–458.

Huang, H. S., & Hanley, R. J. (1994). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*, 73–98.

Küspert, P., & Schneider, W. (1998). *Würzburger Leise Leseprobe [Würzburger Silent Reading Test]*. Hogrefe, Germany: Göttingen.

Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive programme for stimulating phonological awareness in pre-school children. *Reading Research Quarterly, 23*, 163–284.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122–147). Cambridge, MA: MIT Press.

Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading & Writing, 15*, 653–682.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition, 25*, 47–56.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Lawrence Erlbaum.

Muneaux, M., & Ziegler, J. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbor generation task. *Language & Cognitive Processes, 19*, 641–660.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language, 39*, 151–174.

Perfetti, C. A., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly, 33*, 283–319.

Perry, C. (2003). A phoneme–grapheme feedback consistency effect. *Psychonomic Bulletin & Review, 10*, 392–397.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*, 31–74.

Schneider, W., Kuespert, P., Roth, E., Vise, M., & Marx, H. (1997). Short- and long-term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology, 66*, 311–340.

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning & Memory, 5*, 546–554.

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

Slowiaczek, L. M., Soltano, E. G., Wieting, S. J., & Bishop, K. L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology A, 56*, 233–262.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36*, 337–359.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *Journal of Experimental Psychology: General, 124*, 107–136.

Tunmer, W. E., & Nesdale, A. R. (1985). Phonemic segmentation skill and beginning reading. *Journal of Educational Psychology, 77*, 417–427.

Van Orden, G. C., Jansen op de Haar, M. A., & Bosman, A. M. T. (1997). Complex dynamic systems also predict dissociations, but they do not reduce to autonomous components. *Cognitive Neuropsychology, 14*, 131–165.

Ventura, P., Kolinsky, R., Brito-Mendes, C., & Morais, J. (2001). Mental representations of the syllable internal structure are influenced by orthography. *Language & Cognitive Processes, 16*, 131–144.

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language & Cognitive Processes, 19*, 57–95.

Yoon, H.-K., Bolger, D. J., Kwon, O.-S., & Perfetti, C. A. (2002). Subsyllabic units in reading. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (Vol. 11, pp. 139–163). Amsterdam: John Benjamins Publishing Company.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*, 683–689.

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition, 32*, 732–741.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3–29.

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French. *Behavior Research Methods, Instruments, & Computers, 28*, 504–515.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37*, 533–554.

Ziegler, J. C., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language, 48*, 779–793.

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

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, & Computers, 29*, 600–618.