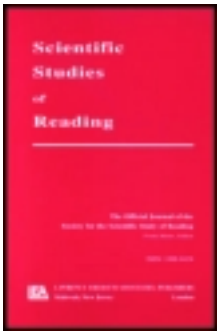

Scientific Studies of Reading
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hssr20

Learning to Read Words in Albanian: A Skill Easily Acquired

Lorenc Hoxhallari , Victor H. P. van Daal & Nick C. Ellis

Published online: 19 Nov 2009.

To cite this article: Lorenc Hoxhallari , Victor H. P. van Daal & Nick C. Ellis (2004) Learning to Read Words in Albanian: A Skill Easily Acquired, Scientific Studies of Reading, 8:2, 153-166, DOI: 10.1207/s1532799xssr0802_3

To link to this article: http://dx.doi.org/10.1207/s1532799xssr0802_3

SCIENTIFIC STUDIES OF READING, *8*(2), 153–166

Learning to Read Words in Albanian: A Skill Easily Acquired

Lorenc Hoxhallari, Victor H. P. van Daal, and Nick C. Ellis

University of Wales

Effects of orthographic transparency were examined by comparing children learning to read in Albanian, Welsh, and English. Twenty Year 1 Albanian children were given a reading test consisting of a 100-word stratified sample of decreasing written frequency. They were able to read accurately 80% of the words; reading latency was a direct effect of word length ($R^2 = .89$); and errors tended to be mispronunciations rather than real word replacements, with hardly any null responses. These results were compared with Ellis and Hooper (2001), where the same design was used with English and Welsh children of the same age, but with 1 more year of formal reading instruction. The Albanian children read more words than the English and Welsh children, but they had longer reading latencies. Like the Welsh children, but unlike the English children, the Albanian children made more nonword errors. These results suggest that children acquire reading faster the more transparent the orthography, and that shallow orthographies promote an initial reliance on a phonological recoding strategy.

Goswami, Gombert, and De Barrera (1998) and Ellis and Hooper (2001) showed that writing system differences between languages may affect the (a) the rate of reading acquisition, (b) young children's awareness of the structure of their spoken language, (c) the sort of reading strategy the children acquire, and (d) the incidence and severity of possible reading disorders. Such cross-linguistic research suggests that existing models of reading, developed primarily for English (e.g., Ehri, 1999; Goswami & Bryant, 1990), may need modification to deal with reading development in shallow orthographies.

With regard to the rate of acquisition, Spencer and Hanley (2003), for example, found that children learning to read in orthographically transparent Welsh were significantly better at reading both words and nonwords than were matched Eng-

Requests for reprints should be sent to Lorenc Hoxhallari, School of Psychology, Brigantia Building, University of Wales, Bangor, Gwynedd LL57 2AS, United Kingdom. E-mail: pspa02@bangor.ac.uk

lish children reading in English. Ellis and Hooper (2001), using materials balanced cross-linguistically with respect to the most important exposure variable of word frequency, also found that Welsh children progressed further in reading than did English children and picked up the alphabetic principle more readily. Thorstad (1991) found that Italian children learn to read in approximately 1 year, whereas English children take 3 to 5 years. However, English children learning to read with the Initial Teaching Alphabet, which was designed to be an orthographically transparent script for English, performed at about the same level as the Italian children. Thorstad suggested that in a transparent orthography where the spelling–sound and sound–spelling correspondences are predictable and invariant, children can rapidly learn these correspondences and then use a systematic, phonological strategy to learn to read and spell.

Wimmer and Goswami (1994) gave groups of 7-, 8-, and 9-year-old German and English children a numeral reading task, a number reading task, and a nonsense word reading task. The nonsense words could be read by analogy to the number words, as these shared parts of the number words; for example: *nive* combines *ni(ne)* and *(fi)ve*. They found that for both orthographies, the reading time and error rates in numeral and number word reading were similar, whereas for the nonsense reading task, even the youngest German children (7-year-olds) made significantly fewer errors than the older English children (9-year-olds). Wimmer and Goswami (1994) suggested that at the beginning of the reading process English children use a larger grain phonological access strategy based on onset or rime, whereas German children begin learning to read by using the finer grained grapheme–phoneme conversions. However, the German children come to use the onset or rime strategy later.

Seymour, Aro, and Erskine (2003) investigated early reading acquisition through assessment of letter knowledge, familiar word reading, and simple nonword reading by measuring accuracy, speed, and error types in 13 different orthographies including English. They found that children from a majority of the European countries became fluent and accurate before the end of the 1st school year. These findings were not present for children learning to read in French, Portuguese, Danish, and particularly English. The rate of reading development in English was more than twice as slow as in more orthographically transparent orthographies. Seymour et al. argued that syllabic complexity and orthographic depth are responsible for the cross-language differences.

This study takes a further step in research into cross-linguistic reading acquisition by examining the acquisition of a completely shallow orthography: Albanian. Other orthographies like Italian, German, Greek, Dutch, French, and Spanish are transparent but only to a lesser extent. Another European script that is nearly completely regular is Finnish, except for one small detail (Lyytinen, Aro, & Holopainen, 2004): the spelling of /ng/ depends on the context. If this sound is followed by a /k/ it is written as *nk*; if not followed by /k/ it is spelled as *ng*. In contrast, Albanian, like Serbo-Croatian, is 100% regular, both from spell-

ing to sound and from sound to spelling. The most frequent syllabic structure in Albanian is Consonant Vowel (CV); CCV clusters are less common.

The writing system is young, hence its transparency (Crystal, 1987). The official alphabet, which was not introduced until 1909, uses Roman characters and has 36 graphemes: 29 consonants, and 7 vowels. Nine of the consonants are presented in the alphabet as bigraphs (*dh*, *gj*, *ll*, *nj*, *rr*, *sh*, *th*, *xh*, *zh*). These bigraphs are explicitly taught to the children as separate graphemes of the alphabet and if they appear together are never considered as separate letters (this rule does not apply to compound words, which occur rather infrequently). The bidirectional grapheme–phoneme and phoneme–grapheme mapping is consistent for both consonants and vowels. Each letter maps to the same sound regardless of whether it is placed initially, medially, or finally, and each sound maps to the same letter regardless of its position in the word.

This study evaluates the reading accuracy of Year 1 Albanian children by comparing their development to those of English and Welsh children examined in Ellis and Hooper (2001). Ellis and Hooper compared the rate of literacy acquisition in Year 2 English and Welsh children matched for reading instruction, background, locale, and math ability. Twenty children from each language group were given a word reading test, where reading accuracy, reading latencies (using a stopwatch), and word error types were assessed. To do this, lists of 100 test-words for English and Welsh were designed by taking the word types that composed million token word frequency counts for English (CELEX database; Baayen, Piepenbrock, & Van Rijn, 1995) and Welsh (Ellis, O'Dochartaigh, Hicks, Morgan, & Laporte, 2001), sorting them in decreasing frequency of occurrence and sampling them so that a test word was selected that matched (roughly) every decreasing step of 10,000 word tokens. Eighty words were selected following this procedure, and another 20 words were randomly selected from words with frequency of 1 per million. Ellis and Hooper argued that once frequency is taken into account, the matching process should not control for other factors such as imageability, word length, utility, morphological complexity, orthographic complexity, and so on, because "everything to do with learning opportunity should be matched; everything to do with language should be freed to vary" (p. 577). Using these tests they determined that (a) Welsh children were able to read aloud accurately significantly more of their language (61% of tokens, 1,821 types) than were English children (52% tokens, 716 types); (b) the reading latencies for Welsh children were longer than for English children; (c) the error patterns differed, with English children producing more visually similar real word substitutions, whereas Welsh children produced more nonword type mispronunciations; and (d) reading latency was a clear function of word length in Welsh but not in English. These findings suggest that Welsh children read by a grapheme–phoneme synthetic strategy, whereas English children rely on lexical retrieval on the basis of partial visual analysis of the written words.

The Welsh alphabet has 29 graphemes, of which 7 are vowels (*a*, *e*, *i*, *o*, *u*, *w*, *y*). The alphabet, like the Albanian alphabet, includes several bigraphs (*dd*, *ff*, *ng*, *ll*, *ph*, *rh*, *th*). The Welsh orthography, unlike English, is highly transparent, enough at least that Welsh dictionaries do not need to illustrate the pronunciation of words using a phonetic notation. Nevertheless, it is not perfectly so: Some written vowels in Welsh map to more than one sound. For example, orthographic *y* may be realized as a schwa in nonfinal syllables of polysyllabic words; it can also represent the first nonvocalic part of a diphthong (*yw*, /iu/) or the second consonantal part of the diphthong (*wy,* /ui/). The Welsh writing system is not as transparent as the Albanian writing system, where each grapheme corresponds only to one sound.

This study seeks to triangulate and replicate the results of Ellis and Hooper (2001) with an Albanian sample. Reading abilities of 20 Year 1 Albanian-speaking children were assessed using a 100-word Albanian test designed to be comparable to English and Welsh lists, and the children's reading was tested using procedures identical to Ellis and Hooper (2001). The first hypothesis is that Year 1 Albanian children will be able to read more words of their language than English children because they use the grapheme–phoneme correspondences that are extremely regular. In comparison to the Welsh children, Albanian children should be able to read more words because the Albanian orthography is even more transparent than the Welsh orthography. It is further expected that the latencies for correctly read Albanian words will be longer than those of the English children because of the extreme transparency of Albanian, leading to these children's exclusive reliance on grapheme–phoneme conversion. Thus, both Welsh and Albanian children should have longer latencies than children reading orthographically opaque English, and Albanian latencies should in turn be longer than those in Welsh, although the magnitude of this latter difference might well be smaller than the differences between either of these languages and English.

Third, it is expected that the errors of the Albanian children will be more like those of Welsh readers in being predominantly non-word-type errors rather than word replacements, because faults in the application of letter–sound decoding result mostly in nonword pronunciation. English children, in contrast, tend to use a strategy of whole-word access on the basis of partial visual analysis or partial phonetic cuing and thus tend to make real word substitution errors.

The children participating in this study were of the same chronological age as the English and Welsh children in the Ellis and Hooper (2001) study. However, it should be noted that most English and Welsh children go to nursery school and play groups at the age of 3 or 4, where they are taught literacy skills such as letter knowledge and sound awareness. Thus, by the time English and Welsh children enter formal education at the age of 5, they already have some knowledge of the letters and the alphabet. In Albania, programs similar to the nursery schools of England and Wales do not exist, and only a minority of preschoolers

go to kindergarten. Many, although not all, of these Albanian kindergartens are similar in ethos to the Austrian ones described by Wimmer and Hummer (1990), where activities such as word exposure and letter recognition are explicitly discouraged. The consequence is that most first graders in Albania enter school without letter or written word awareness (Hoxhallari, 2000). Thus our three language groups are matched for chronological age, but the Albanian children have had less print exposure and less reading instruction. These aspects of the design thus bias against any literacy advantage in the Albanian group. Similarly, although we did not assess for socioeconomic status, Albania is a much less wealthy country than Britain, with concomitantly fewer resources to invest in education and health. This is another factor that would bias against our first and second hypotheses.

METHOD

Participants

Twenty monolingual Albanian children from a primary school located in a city in southeast Albania took part in the experiment. The participants were medium-achieving pupils on the most recent math test administered by the classroom teacher—a method, employed by Bast and Reitsma (1997), to exclude the worst and best students from the study. Although this selection procedure differed from that of Ellis and Hooper (2001), where whole classes were examined, the results are still comparable as we look at average achievements and not the range of individual differences. Despite the fact that the Albanian orthography is extremely transparent and used to be taught with a phonics method, the Albanian government has recently prescribed a whole-word method, known as the global method, for all state schools in the country. Therefore many experienced teachers, including the teacher who taught the children who participated in this research, use a mix of the global/whole-word and phonics method (introducing grapheme–phoneme correspondences from the beginning of the school year). Thus, the teaching method used to teach literacy to the Albanian children who participated in this research is similar to teaching methods employed in English and Welsh schools in that it is a mixture of techniques. Again, the current predominance of Albanian national curriculum should bias against our hypothesis that this transparent orthography will promote an alphabetic rather than a whole-word reading strategy.

There were 8 boys and 12 girls, with a group mean age of 7 years 6 months (*SD* = 6 months). According to the classroom teacher, all the participants were normally developing children with no reading or spelling difficulties, and they came from a mix of social backgrounds. There is no reason to believe that the school attended by the children was either a very good or a very poor performing school.

We believe therefore that the participants are a representative sample of Albanian school children. Albanian children had received 8 months of formal reading instruction prior to testing in the second semester of the school year. The participants in the Ellis and Hooper (2001) study were between 6 and 7 years old and attended Year 2. Altogether there were 17 girls and 23 boys, 20 from a Welsh primary school and 20 from an English primary school, matched with respect to exposure to print.

Reading Test

The Albanian reading test was designed by randomly sampling one word from each of 100 successive strata of decreasing log10 written word frequency from an Albanian text corpus created by the first author, comprising just over 1 million words, from a novel, the New Testament, one children's book, one online Albanian newspaper (dates ranging form early January to late February 2001), and two short passages from books advertised on the Internet. This method is broadly equivalent to the 1-in-10,000 token sampling method used by Ellis and Hooper (2001) but slightly improves on it in that it produces strata that are separated by equal intervals of log10 frequency. The most frequent word of the list was *të* 'you' with a frequency of 67,713 in 1 million, and one of the last words was *lëvdoni* 'praise', which occurred only once. Any effects of the slight differences in frequency profile of reading tests resultant from the variation in sampling method across the two studies will be checked in following analyses and controlled by analysis of covariance (ANCOVA).

Table 1 shows the frequency distribution of the length of the words for the three languages and shows that the lists are well matched in their sampling in word length (Albanian $M = 5.36$, $SD = 2.94$; English $M = 5.84$, $SD = 3.14$; Welsh $M = 5.60$, $SD = 2.97$), $F(2, 98) = 1.47$, *ns.* For the subsequent analyses, words containing less than four letters are regarded as short words, whereas words containing five letters or more are regarded as long words. There are 51 short words and 49 long words.

The word list was printed with 17 words per A4-size page, double spaced, centered on the line and set in bold, lowercase, 20-point Times font. Accuracy and reading latency (for correct answers) were measured. The maximum score for reading accuracy was 100.

Procedure

Each participant was tested by a pair of experimenters. Three pairs of experimenters worked in one large room, each pair far from another so that the participants

TABLE 1
Word Length and the Frequency for the Albanian, English, and Welsh Reading Test

Word Length (No. of Letters)	*Frequency*		
	Albanian	*English*	*Welsh*
1	4	2	5
2	15	12	11
3	11	12	12
4	18	14	9
5	10	16	16
6	10	7	11
7	8	10	13
8	7	5	7
9	6	6	7
10	5	6	4
11	2	5	2
12	3	2	0
13	1	2	1
14	0	0	1
15	0	1	0
16	0	0	1

could not disturb each other. The word reading procedure was identical to the Ellis and Hooper (2001) study. The participants were asked to read the words and to speak loudly and clearly because they would be recorded on a tape. The tape was used as a backup in case the experimenter missed the timing of the response. A piece of plain card was used to cover the list, and the child was asked to move the card down when the first experimenter said "Next," and to read the following word immediately. A stopwatch was used to record latencies from the onset of word presentation to the response onset. Using a stopwatch like this provides sufficient reliability to clearly discriminate individual differences between readers and between words; for example, the interobserver correlation between the reaction times measured by two separate experimenters measuring the word readings of the same child was highly significant ($r = .93$, $p < .01$). If the child made no response after 15 sec, the experimenter would ask the child to try the next word. Each child was given up to four practice trials to show that they understood the instructions. The reading test was discontinued if five consecutive errors were made. The reading errors were written down on a separate sheet by the second experimenter.

RESULTS

Group differences were assessed using by-item analyses of variance.

Reading Accuracy

The Albanian children read on average up to the 80th word from the reading list (*SD* = 16.42). In the Ellis and Hooper (2001) study the English children read on average up to the 52nd word (*SD* = 11.7) and the Welsh children read up to the 61st word (*SD* = 16.2). These group differences are significant, $F(2, 98) = 40.4$, $p < .001$. Post hoc Bonferroni tests showed that the Albanian children were significantly better at reading accuracy than both the English and the Welsh children.

To verify that these effects are not due to minor differences between the token frequency distributions of the parallel test items in the three languages, we also assessed these effects of language on accuracy using ANCOVAs with log10 frequency as the covariate. The language group differences remained highly significant, $F(2, 98) = 7.84$, $p < .001$.

For the Albanian sample, gender ratio had no effect on reading accuracy (by subject analysis).

Reading Latency

Albanian children read at an average pace of 2.66 sec/word (*SD* = 1.58), with the short words being read at an average pace of 1.49 sec/word (*SD* = 0.65) and long words at 3.87 sec/word (*SD* = 1.33). The English children read at an average pace of 1.41 sec/word (*SD* = 1.13), and the Welsh children at 1.85 sec/word (*SD* = 1.11). These group differences are significant, $F(2, 98) = 16.8$, $p < .001$. Post hoc Bonferroni comparisons showed that the Albanian children's reading latencies were significantly longer than both English and Welsh children's reading latencies. These effects of language group on latency remain significant when tested by ANCOVA with log10 frequency as the covariate, $F(2, 98) = 3.94$, $p < .05$.

Regression analysis for the Albanian group relating latency of correct responding to log10 frequency shows that the reading latency increases by 0.89 units for each decreasing log10 frequency step ($B = –0.89$, *SE* B = 0.079, $\beta = –0.75$), $F(1, 98) = 124$, $p < .01$. The log10 scale for frequency makes the relation between frequency and latency more linear. The increase in latency of 0.89 units for each decreasing frequency step means that it takes 0.89 seconds more to read a word out of each 10 successive strata of decreasing log10 written word frequency.

Figure 1 depicts the reading latencies for correctly read items in Albanian, Welsh, and English. It is clear that for the first 30 to 40 items the latencies for all

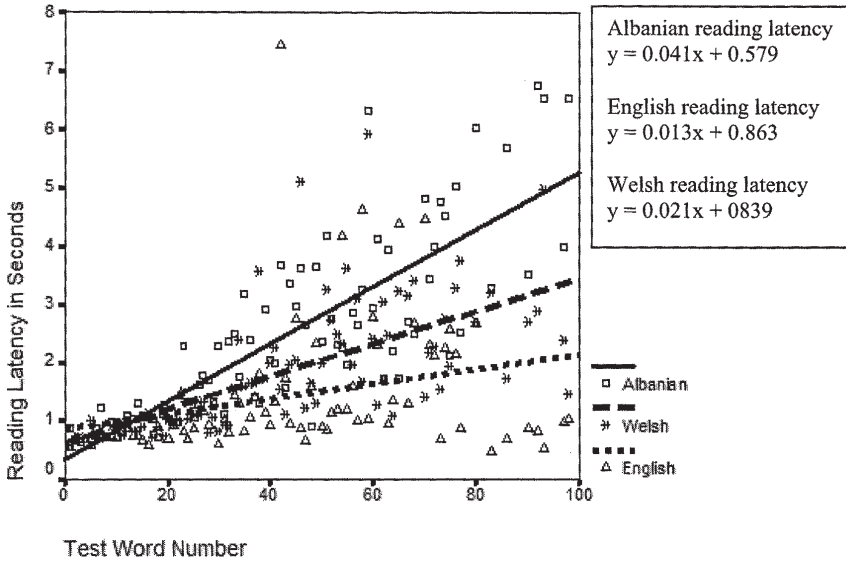

FIGURE 1 Reading latencies for correctly read words in Albanian, English, and Welsh.

groups are not very different. The main differences in word latencies appear to be for the last 50 items, especially between the English and Albanian children. A regression of reading latencies on test word number resulted in the following functions: for the Albanian children, $y = 0.041x + 0.579$; English children, $y = 0.013x + 0.863$; and Welsh children, $y = 0.021x + 0839$; Albanian $B = 0.041$, SE B = 0.004, $\beta = 0.75$, $F(1, 98) = 128$, $p < .0001$, $R^2 = 0.56$; English $B = 0.013$, SE B = 0.005, $\beta = 0.30$, $F(1, 98) = 7.71$, $p < .01$, $R^2 = 0.087$; Welsh $B = 0.021$, SE B = 0.03, $\beta = 0.58$, $F(1, 98) = 48.61$, $p < .001$, $R^2 = 0.33$. Using the z statistic (Sachs, 1982) it was found that the R^2s differ significantly from each other across all three language groups (all zs > 1.96, $p < .05$).

Reading Latencies as a Function of Word Length

Regression analysis reveals that for the Albanian children, reading latency is a function of word length: $B = –0.50$, SE B = 0.021, $\beta = –0.93$, $F(1, 98) = 587$, $p < .001$. Figure 2 shows that reading latency for correct responses is more a linear increasing function of word length for the Albanian children ($R^2 = .86$) than for the

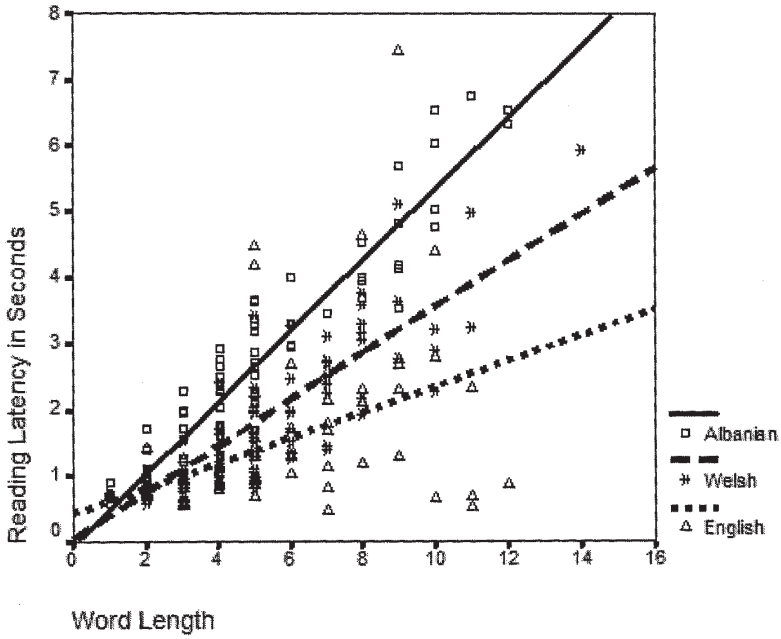

FIGURE 2 Latency as function of word length for words correctly read aloud.

English (R^2 = .22) and the Welsh children (R^2 = .68). Using the z statistic, it was found that the R^2s differ significantly from each other across all three language groups (all zs > 1.96, $p < .05$).

Reading Errors

In total, 296 errors were made by the Albanian children. The errors shown in Table 2 were classified as null responses (no response within 15 sec), word replacements, or other attempts that resulted in nonword responses. Most of the errors made by the Albanian children were nonword errors (64.9%), whereas null responses were very rare (only 4.4%). Whole-word substitutions (30.7%) were not as frequent as nonword errors. A chi-square test shows that there is a significant pattern in the error distribution, $\chi^2(2, N = 296) = 163.26$, $p < .001$, for the Albanian children.

Table 2 also shows the distributions of Welsh and English number of errors for comparison. There is a significant association between error types and language groups, $\chi^2(4, N = 1{,}039) = 166.67$, $p < .001$. The Albanian and Welsh error patterns are very similar to each other, whereas the English errors are quite differently distributed. Haberman's residuals highlight particularly fewer null responses for Al-

TABLE 2
Frequency of Error Types for Albanian, Welsh, and English Children

	Albanian		*Welsh*		*English*	
Error Type	*n*	*%*	*n*	*%*	*n*	*%*
Null responses	13	4.4[a]	52	13.7	112	30.8[a]
Word replacements	91	30.7	92	24.3[a]	163	44.8[a]
Nonword errors	192	64.9[a]	235	62.0[a]	89	24.5[a]
Total	296		379		364	

[a]Haberman Standardized Residuals > 2.

banian children (4.4%) than for English (30.8%), that the Albanian and Welsh children commit more nonword errors (64.9% and 62.0%, respectively) than the English children (24.5%) and that the English children tend instead to make more real word replacement errors.

DISCUSSION

These data triangulate and confirm the conclusions of Ellis and Hooper (2001) with respect to the effect of orthographic transparency on the rate of reading acquisition and the strategies children use in the beginning stages of learning to read.

The rate of reading acquisition is faster the shallower the orthography: The Albanian children were able to read correctly more words than the Welsh children and far more words than the English children of Ellis and Hooper (2001), with Albanian children reading approximately 80% of the reading test items of their language after just 1 year of school education. This percentage covers the 8,021 most frequent word types of Albanian (equivalent to approximately 87% of written Albanian tokens), whereas from the Ellis and Hooper study, English children could read only about 52% of the English test items, a coverage generated by the 716 most frequent word types (approximately 64% of written English tokens). Comparable Welsh children read approximately 61% of their test, a coverage generated by the 1,821 most frequent word types of Welsh (approximately 81% of tokens). This finding is consistent with other studies of the effect of transparency on rate of acquisition, including Wimmer and Goswami (1994), Öney and Durgunoğlu (1997), Seymour et al. (2003), and Spencer and Hanley (2003).

Two strands of evidence show that beginning readers in the shallow orthography of Albanian rely on a grapheme–phoneme conversion strategy: (a) the strong effect of word length on their reading latency and (b) their high proportion of nonword errors.

Albanian children's reading latencies were greatly affected by word length. This is consistent with their attempting to pronounce long, novel, or infrequent words by phonological recoding using a left-to-right parse; thus, the more letters in a row, the longer the production time. This finding is reaffirmed by the fact that Albanian children's latencies were longer than English children's and longer even than the Welsh children's. This suggests that the greater the orthographic transparency, the more likely a phonological reading strategy is to be consistently employed. One might have expected that Albanian and Welsh would be more or less similar in this respect because both orthographies are highly transparent, yet the complete transparency of Albanian ties children's reading latencies even more tightly to word length.

The second strand of evidence comes from reading error analyses. Unlike their English counterparts, the Albanian children made more nonword errors than any other kind of error. This high proportion of nonword errors suggests that their reading strategy relies heavily on alphabetic recoding, as demonstrated for reading of other transparent scripts by Wimmer and Hummer (1990), Spencer and Hanley (2003), Seymour et al. (2003), and Ellis and Hooper (2001). When phonological recoding goes wrong, the likely pronunciation error is a nonword; when whole-word lexical access on the basis of partial visual analysis goes wrong or when partial phonetic cuing is employed, the likely result is a word. A phonological reading strategy also means that children are prepared to have a go at reading any word: Despite the fact that the Albanian children have had less time on task reading than the English children, they tended to make very few null responses. Regularity of orthography gives children greater confidence in trying to read new words. Following this strategy they will develop a self-teaching mechanism, which allows them to enlarge their sight-word reading lexicon (Share, 1995).

Considering latency and error data together we can say that Albanian children went further in the lists and encountered relatively longer words that took them more time to read, because they continued to apply a phonological recoding strategy. They made fewer errors, and if they made an error, nonword substitutions prevailed. On the other hand, both English and Welsh children made more mistakes and gave more null responses, and English children replaced words with other words. The Albanian children were not slower than the English and Welsh children as a result of less reading experience; if this were the case, we would have expected them to be slower at the first 40 word-items, but they were not. Here the steepness of the functions is relevant: In Albanian children there is a steeper word length effect and a steeper frequency effect than in the children of the other two languages. Because frequency and word length are highly correlated ($r = -.75$ for each language), it is likely that indeed a phonological recoding strategy was used by the Albanian children.

In conclusion, children learn to read an orthographically transparent script very quickly. Their reading accuracy is high. They quickly adopt a phonological reading strategy that ties reading latency to word length, that generates errors that tend

to be nonword pronunciations, and that gives these children confidence in having a go at new words.

ACKNOWLEDGMENTS

This work was supported by the Economic and Social Research Council (ESRC award H474 25 4101).

We are grateful to the teachers, children, and the head of Sotir Gurra School, Korçë, Albania, for all their help. We thank Marilyn Vihman and Heikki Lyytinen for their explanations of the Finnish script.

REFERENCES

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). *The Celex lexical database release 2.* Nijmegen, Netherlands: Centre for Lexical Information, Max Planck Institute for Psycholinguistics.

Bast, J., & Reitsma, P. (1997). Matthew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32,* 135–167.

Crystal, D. (1987). *The Cambridge encyclopedia of language.* Cambridge, England: Cambridge University Press.

Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading* (pp. 79–108). Oxford, England: Blackwell.

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

Ellis, N. C., O'Dochartaigh, C., Hicks, W., Morgan, M., & Laporte, N. (2001). *Cronfa electroneg o gymraeg* [A 1 million word lexical database and frequency count for Welsh]. Bangor, UK: University of Wales, Bangor. Retrieved August 20, 2001, from http://www.bangor.ac.uk/ar/cb/ceg/ceg_eng.html

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

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. *Journal of Applied Psycholinguistics, 19,* 153–162.

Hoxhallari, L. (2000). *Teacher's beliefs about learning to read Albanian, English and Welsh: The effect of a regular orthography.* Unpublished manuscript, University of Wales, Bangor.

Lyytinen, H., Aro, M., & Holopainen, L. (2004). Dyslexia in highly orthographically regular Finnish. In I. Smythe, J. Everatt, & R. Salter (Eds.), *International book of dyslexia* (2nd ed., pp. 81–91). London: Wiley.

Öney, B., & Durgunoğlu, A. Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics, 18,* 1–15.

Sachs, L. (1982). *Applied statistics* (5th ed.). New York: Springer-Verlag.

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

Share, D. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Spencer, L. H., & Hanley, R. J. (2003). Effects of orthographic transparency on reading and phoneme awareness in children learning to read in Wales. *British Journal of Psychology, 94,* 1–28.

Thorstad, G. (1991). The effect of orthography and the acquisition of literacy skills. *British Journal of Psychology, 82,* 527–537.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51,* 19–103.

Wimmer, H., & Hummer, P. (1990). How German speaking first graders read and spell: Doubts on the importance of the logographic stage. *Applied Psycholinguistics, 11,* 349–368.

Manuscript received June 17, 2002
Accepted June 14, 2003