Journal of Educational Psychology
1984, Vol. 76, No. 4, 557–568

Decoding and Comprehension Skills in Turkish and English: Effects of the Regularity of Grapheme–Phoneme Correspondences

Banu Oney and Susan R. Goldman
University of California, Santa Barbara

The purpose of this study was to examine acquisition of reading in languages with different degrees of grapheme–phoneme correspondence. Decoding and comprehension skills were assessed for Turkish and American first and third graders learning to read Turkish or English, respectively. Twenty students in each group were tested on a pseudoword vocalization task and on a paragraph comprehension task. Latency and accuracy data indexed decoding proficiency and percentage correct comprehension skill. Results generally confirmed predictions. Turkish students were faster and more accurate on the decoding task than Americans at the first-grade level and equally accurate but faster at the third-grade level. In first grade, but not third, Turkish students were superior to Americans on the comprehension task. Significant relationships between decoding and comprehension were found for all but American third graders. The data suggest that languages with more regular letter-sound correspondences lead to faster acquisition of decoding skills. The findings also support a decreased relationship between decoding and comprehension once learners are beyond initial reading.

Several researchers have suggested that the acquisition of reading is affected by the nature of the orthography of the language that the child is learning to read (Kyostio, 1980; Malmquist & Grundin, 1980; Rozin & Gleitman, 1977; Venezky, 1972). Although few languages show perfect one-to-one correspondence between letters and sounds, some languages manifest greater regularity than others. Liberman and Shankweiler (1979) cited Serbo-Croatian languages, such as Turkish and Hungarian, and Finnish as examples of alphabetic languages in which the mapping from written symbols to phonemes is more nearly one-to-one than in the English language. Although English orthography does follow a set of patterns and rules (Venezky, 1967, 1970), these regularities tend to be found in units consisting of more than one letter. Furthermore, spelling and pronunciation patterns are often contextually determined. Such irregularity and variability has been cited as a factor that makes learning to read English more difficult than learning to read languages with more regular and predictable grapheme-phoneme correspondences (Liberman & Shankweiler, 1979).

The greater difficulty associated with irregular letter–sound correspondences might be expected to have its greatest effect on learning to "crack the code," with decreasing importance as higher order reading skills are acquired. For example, Kyostio (1980) and Venezky (1972) have reported high degrees of letter-sound mastery for children learning to read Finnish and only moderate correlations between letter-sound skills and standardized measures of reading ability. Such a finding is hardly surprising from either a statistical or theoretical perspective. For example, Chall (1979), as well as a number of others (e.g., LaBerge & Samuels, 1974; Perfetti & Lesgold, 1979), pointed out the importance of conceptualizing reading as a skill comprising a number of subskills. At different phases in reading acquisition, certain subskills are more focal than others. Initial reading skill is heavily dependent on

This research was conducted as Banu Oney's master's thesis, Graduate School of Education, University of California, Santa Barbara.

The valuable contributions of committee members Priscilla Drum and James W. Pellegrino are acknowledged.

Requests for reprints should be sent to Susan R. Goldman, Graduate School of Education, University of California, Santa Barbara, California 93106.

decoding (i.e., on mastery of the letter–sound mappings). Later, reading requires skills associated with connecting and integrating the meanings within and across units expressed in the written medium.

However, the interactive nature of subskills in reading, coupled with limited processing capacities, has led a number of researchers to suggest that facility at decoding affects higher order comprehension skills. LaBerge and Samuels (1974) argued that one of the important aspects of reading acquisition involves the reduction of attentional resources demanded by lower levels of processing. It is assumed that automaticity of preliminary processing leads to the availability of attentional resources for higher order processing such as comprehension. Studies concerned with individual differences in reading and the effects of decoding speed on comprehension suggest that due to the sharing of a limited-capacity mechanism, readers slow at decoding individual words have less capacity available for higher order processes (e.g., Curtis, 1980; Lesgold & Perfetti, 1981; Perfetti, 1977; Stanovich, 1981). A study conducted by Curtis (1980) emphasizes both the relationship between decoding and comprehension and changes in the importance of decoding speed over the time course of reading acquisition. Curtis found that the speed of processing printed material significantly increased at successive grade levels. However, there was also a decrease in the amount of variance in reading ability explained by the reading speed measures. Thus, at initial stages of reading acquisition, regularity of grapheme–phoneme correspondences exerts a greater effect on comprehension than during later stages. As readers learning to read in languages with differing degrees of letter-sound correspondence advance in reading skill, they would be expected to develop automaticity in word recognition, which in turn would decrease the contribution of decoding to reading skill.

The present study examined the general question of whether it is easier to learn to read in a language that has more regular letter–sound correspondence by comparing Turkish students learning to read Turkish with American students learning to read English. Turkish and English were selected because they are representative of languages with different degrees of grapheme–phoneme correspondence, with Turkish having higher regularity than English. Both languages use English orthography, Turkey having adopted this alphabet in the mid 1920s. Consistent with the suggestion of Malmquist and Grundin (1980), we hypothesized that beginning readers of Turkish would make faster progress than those of English. We expected this to be manifested by faster and more accurate decoding of pseudowords. We also addressed the issue of the decreasing importance of facile decoding in reading by testing first and third graders on a paragraph comprehension task as well as the decoding task. Thus, we examined the relationship between decoding proficiency with pseudowords and comprehension performance in students just learning to read (first grade) and in a more advanced group of readers (third grade). The specific hypotheses tested were as follows: (a) Turkish readers would be more proficient at decoding than Americans in the first grade, but this difference would decrease by the third grade. This prediction was based on the notion that decoding automaticity would take longer to develop for students learning English than for students learning Turkish. (b) Comprehension skills of Turkish first graders would be superior to American first graders, but no differences were expected in third grade. This prediction was based on the notion that by the first grade, students learning to read Turkish would have developed decoding skills sufficient to enable them to comprehend discourse, whereas for those learning English, it would take a longer period of time to reach this level. By third grade, we expected that differences in decoding skill proficiency would not be contributing to comprehension skills in the same manner. (c) The relationship between decoding proficiency and comprehension skills would decrease within each language group from first to third grade.

Method

Subjects

Subjects were selected from one Turkish and one American elementary school in Turkey, attended by native Turkish and American students, respectively. Twenty subjects were selected from the first and from the third grades at each elementary school by having the classroom teachers rank order the students on the basis of their reading ability. Five students who were ranked lowest in each class were not included in the sample. In addition, students who were repeating the same grade were excluded from the sample. Mean chronological age was 7.2 for the American first graders, 7.4 for the Turkish first graders, 9.5 for the American third graders, and 9.2 for the Turkish third graders. No information could be obtained on the results of reading readiness or reading achievement tests because such tests are not used in Turkish schools, and those that were administered to the American students were not available to the experimenter. All testing was conducted during May of the school year.

The reading programs in each school were similar. Both used phonics and basal reader materials. Whereas 10 years ago the Turkish schools tended to have a greater emphasis on phonics than did the American, this is no longer the case. The American school from which subjects were sampled is located on a military base. The children attending the Turkish school represent a relatively equivalent socioeconomic status (SES) level: Most of their parents are employed in lower level government bureaucracy positions. Class size in each school was between 20 to 25 pupils per class.

Materials

A list of 84 pseudowords, pronounceable in both Turkish and English, was generated by changing the initial consonant in each syllable of 42 real words from each language. The list consisted of 21 one-syllable, 36 two-syllable, and 24 three-syllable pseudowords.[1] The real words from which the pseudowords were generated were selected on the basis of the four syllable patterns, and their combinations, common to Turkish and English. In changing the initial consonants of the syllables, patterns not permissible in both languages were avoided. In the case of consonant clusters, the clusters were replaced by ones permissible in both languages. When the syllable began with a vowel instead of a consonant, the final consonant in the syllable was changed. The complete list of pseudowords is shown in Appendix A. Note that whereas Turkish sometimes uses an umlaut (¨), the absence of an umlaut in the pseudoword materials merely changes the correct pronunciation of the specific letter string.

Another list of 24 one-syllable pseudowords was constructed from syllable patterns and three-letter consonant clusters that occur in English but not in Turkish. Pseudowords were constructed by changing the initial consonant or by replacing 1 three-letter consonant cluster with another. The list of pseudowords constructed from these nonoverlapping syllable patterns was presented to both the Turkish and American subjects, but these were treated as filler words and were excluded from the analyses for Turkish subjects. The rationale for presenting the list of words with nonoverlapping syllable patterns to the American subjects was to see if the syllable patterns limited to English, such as those containing consonant clusters of more than two consonants, present any particular difficulty to the American subjects. These patterns essentially typify the irregular character of letter–sound correspondences in English. The data on these pseudowords from the Turkish subjects were not analyzed because there are no correct pronunciations for these letter strings in Turkish.

In order to rule out the possibility of item position effects, words were arranged in 12 blocks of 9 words. Each block consisted of 2 nonoverlap, 2 one-syllable, 3 two-syllable, and 2 three-syllable words. In 6 of the 12 blocks, of the 7 overlap pseudowords, 4 were derived from Turkish and 3 from English; in the other 6 blocks 4 were derived from English and 3 from Turkish. The 2 nonoverlap pseudowords were derived from English. Words were assigned randomly to blocks in such a way that not more than two of each type appeared in succession. The blocks were further arranged in two orders, (1–12, 12–1), with half of the subjects in each grade by nationality group receiving one order and the other half receiving the other order.

Acceptable pronunciations for each word were obtained from two linguists in both nationalities in order to determine those pronunciations to be considered accurate.

In order to assess comprehension skills, two paragraph-length texts were selected from the primary and elementary forms of the Metropolitan Achievement Test. The primary level is typically administered to first and second graders, the elementary level to third and fourth graders. The texts were then translated into Turkish, trying to maintain the same readability levels. The content selected was judged to be appropriate for Turkish students by the first author and two other Turkish graduate students. Umlauts were, of course, used where appropriate. Five literal questions relating to each story were also prepared and used to assess comprehension. The English versions of the texts and questions are given in Appendix B.

Procedure

Subjects were presented with the pseudoword vocalization task followed by the comprehension task. During the pseudoword vocalization task, words were projected on a screen one by one, and subjects were instructed to read them as quickly and as accurately as possible. The image of the word automatically disappeared as soon as the subject started to vocalize the word. This was accomplished through the use of a Lafayette voice-operated relay that was connected to the shutter on the slide projector. A millisecond timer was activated by the opening of the shutter and was stopped by the voice-operated relay. Vocalization latencies and accuracy of production were recorded for each word.

In order to acquaint the subjects with this task, a practice session was held before starting the task, where

[1] The original list of one-syllable pseudowords was to have contained six examples of each of the four patterns. Due to experimenter error, one pattern (vowel/consonant/consonant; VCC) contained only three examples.

12 words were used as practice words. The first two practice words were different for Turkish and American subjects and were easy real words from their native language. The remaining 10 words were pseudowords generated from real words in the same fashion as the *experimental pseudowords. Subjects were told to* pronounce the word only when they were ready to say the whole word because the image of the word would disappear with the onset of vocalization. None of the subjects had any problems with this procedure.

The comprehension task was presented after the vocalization task. All subjects read both the primary- and elementary-level stories. They were asked to read the stories orally. After reading one, they were asked to answer the five questions, which were presented to them one by one orally by the experimenter. Subjects received a score ranging from 0 to 5 on each paragraph, based on the number of questions answered correctly. The order of presentation of the texts was counterbalanced, so that within each group half of the subjects were presented with the primary-level text first and half were presented with the elementary-level text first.

Results and Discussion

Vocalization Latency

The number of syllables in the pseudowords was to be the primary unit of analysis for the decoding data. However, there were different consonant–vowel patterns and numbers of letters within the syllable types. Therefore, in order to examine whether the patterns within syllable types had an effect, 14 syllable types were analyzed. This analysis led to the realization that syllable length accounted for about one third of the between-item sums of squares, and letter length within syllable type accounted for the remainder. A second analysis was done on nine syllable types representing number of syllables and letter lengths within syllable types. An examination of these data indicated that patterns of same letter length but different number of syllables did not differ. Therefore the 14 patterns were collapsed into seven types, based on letter length, and the primary unit of analysis was changed from the number of syllables to letter length. The results of these preliminary analyses are available from the second author upon request.

Analysis of variance (ANOVA) of the mean latencies for correctly pronouneed pseudowords was conducted with letter length (three to nine) as the within-subjects variable.[2] Between-subjects factors were nationality (Turkish and English) and grade (first and third). The ANOVA design was therefore a 2 × 2 × 7, with repeated measures on the letter-length factor. The results indicated that the American subjects (M = 3.233) had longer latencies than the Turkish (M = 1.816), $F(1, 69) = 43.33$.[3] First graders (M = 3.411) took longer than third graders (M = 1.639), $F(1, 69) = 68.99$. The Nationality × Grade interaction was nonsignificant ($F < 1$). There was also a significant effect for letter length, $F(6, 449) = 81.94$. The linear trend component of this main effect was significant, $F(1, 449) = 466.36$, and accounted for 95% of the sums of squares due to this factor. The means for each of the seven letter lengths are given in the last column of Table 1. There were also significant first-order interactions of Nationality × Letter Length, $F(6, 449) = 9.46$, and Grade × Letter Length, $F(6, 449) = 6.95$, and a significant second-order interaction of Nationality × Grade × Letter Length, $F(6, 449)$

Table 1
Mean Vocalization Latencies on Seven Letter Lengths

No. of letters	Grade: American 1	American 3	Turkish 1	Turkish 3	Across nationalities and grades
3	2.459	1.250	1.668	0.764	1.535
4	3.570	1.606	1.857	0.862	1.974
5	4.233	1.888	2.403	0.939	2.366
6	4.036	1.926	2.574	0.962	2.375
7	4.858	2.800	3.088	1.139	2.971
8	5.355	3.131	3.173	1.236	3.224
9	4.952	3.200	3.522	1.238	3.228

[2] As discussed by Pachella (1974), mean response time is preferable to a measure that would arbitrarily eliminate the effects of outlying data points. Contributions of outlying points should be minimal due to the large number of data points available for each subject.

[3] Unless otherwise indicated, statistical significance levels are $p < .01$. Note further that the degrees of freedom for the within-subjects latency data have been adjusted to compensate for seven data points that had to be estimated because 6 subjects had no correct pronunciations in these cells. Five of these estimates were made for American first graders on Letter-Length 9 and two of these estimates were made for 1 American third grader on Letter-Length 7 and on Letter-Length 9. Estimates were based on the average of the subject's mean and the mean for the particular letter string.

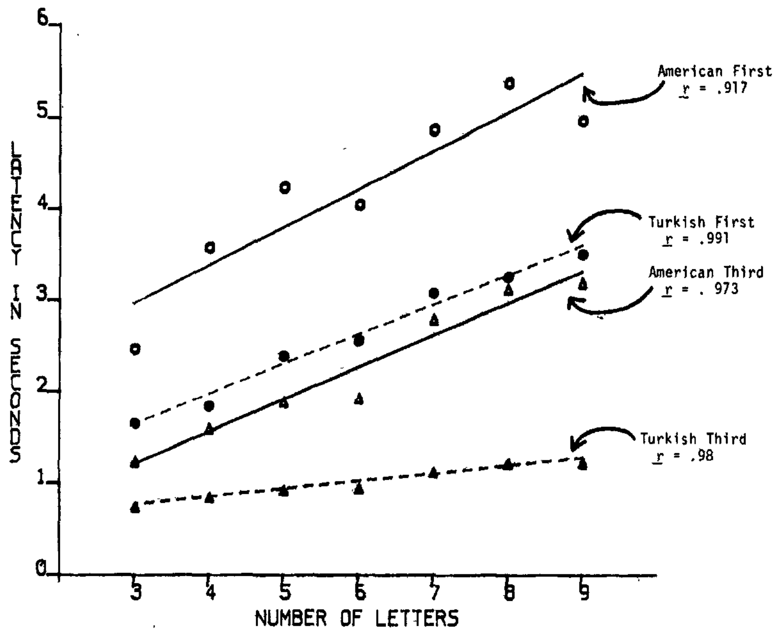

Figure 1. Relation between mean vocalization latency and the number of letters in the pseudowords for each grade and nationality group.

= 3.12. The respective cell means are also given in Table 1.

The difference in the linear trend component of the Nationality × Letter Length interaction was significant, $F(1, 449) = 44.18$, and accounted for 78% of the sums of squares due to this interaction. The value of the linear trend for the Americans was greater than that for the Turkish subjects. The difference in the linear trend for the Grade × Letter Length interaction was also significant, $F(1, 449) = 32.4$, and accounted for 78% of the sums of squares due to this interaction. The value of the linear trend for the first graders was greater than that for the third graders. Finally, the difference in the linear trend component for the Nationality × Grade × Letter Length interaction was significant, $F(1, 449) = 9.36$, and accounted for 50% of the sums of squares due to this interaction. The linear trends within each nationality by grade group are shown in Figure 1. Simple effects tests confirmed that the linear trend was significant in each of the groups, all Fs$(1, 449) > 4.0$, $p < .05$. Similarly, the least squares regression lines shown in Figure 1 reflect significant correlations between letter length and mean vocalization latency within each of the groups of subjects (all rs $> .91$, $df = 18$). These least squares regression lines were fit to the group mean data.

The significance of the linear trend in the main effect of letter length and within each of the four subject groups, in combination with the significance of the difference in the linear trend components in accounting for the sums of squares in the various interactions, suggested the importance of examining the slope and intercept data on an individual subject basis. Thus correlation-regression analyses were applied to each subject's data. Slope and intercept measures were computed, based on the subject's mean latency score on each of the seven letter lengths.[4] These slopes and intercept values were then submitted to ANOVAs.

The analysis of the intercept data indicated that the intercept was higher for the first graders ($M = 1.14$) than for the third graders ($M = 0.34$), $F(1, 76) = 12.61$, and

---

[4] In computing the slope and intercept for each subject, if a subject was missing a data point for one of the letter lengths, the calculation was based on the remaining data points. Except for 1 American third grader, no subject's slope or intercept was based on less than six means. That subject's data were based on five means.

that there was no effect of nationality, $F(1, 76) = 1.54$, mean intercept = 0.7. There was, however, a significant Grade × Nationality interaction, $F(1, 76) = 7.87$. Simple main effects tests indicated that first-grade Americans had a higher intercept value ($M = 1.6$) than American third graders ($M = 0.16$), $F(1, 76) = 20.08$; however, the first- ($M = 0.68$) and third-grade ($M = 0.51$) Turkish subjects did not differ significantly ($F < 1$).[5] The intercept measure may reflect a perceptual-motor response factor, perhaps related to task orientation. It is interesting to note, although difficult to interpret, the grade difference on this measure in American but not in Turkish subjects.

The analysis of the slope data indicated that Turkish subjects had a shallower slope ($M = 210$ ms) than the American subjects ($M = 400$ ms), $F(1, 76) = 21.55$. The third graders had a shallower slope ($M = 220$ ms) than the first graders ($M = 380$ ms), $F(1, 76) = 16.06$. Consistent with prior analyses, there was a Grade × Nationality interaction ($F = 3.35$), although this interaction exceeded conventional levels of statistical significance ($p < .07$). These slope differences are consistent with our hypotheses. The slope reflects the cost per additional letter. As can be seen in Figure 1, the cost per additional letter decreased from a mean of 439 ms for the American first graders to 350 ms for the American third graders. A decrease in processing time per letter was also observed between Turkish first graders (324 ms) and Turkish third graders (85 ms). The processing time for each additional letter may be compared to Frederiksen's (1978) findings. In a study involving high- and low-ability readers from the high school level, he found that the processing time per additional letter was about 45 ms for low-ability readers and 20 ms for high-ability readers on pseudowords. He also found that the cost of each additional letter decreased when real words were compared to pseudowords, and high-frequency words were compared to low-frequency words (Frederiksen, 1978, 1980). Based on these findings, differences in the cost per additional letter might be expected to be smaller with real words.

With respect to decoding speed, the data support the hypothesis that individuals learning to read in Turkish, which has a higher degree of letter-sound correspondence than English, are faster in decoding: Turkish subjects at both grade levels were faster than American subjects at the same grade levels. In addition, decoding speed improved between first and third grades in both nationalities, consistent with the development of greater automaticity at this level of processing in reading.

Decoding Accuracy

It could be argued that the vocalization latency results reflect a speed–accuracy trade-off: The Turkish subjects showed a shallower slope because they made more errors. This was not the case. For each subject the percentage of accurately pronounced pseudowords on each of the seven letter lengths was computed. An ANOVA conducted on the percentages indicated that the Turkish subjects were more accurate (94.6%) than the Americans (73.2%), $F(1, 76) = 60.52$, and that there was an effect of letter length, $F(6, 456) = 34.25$. As the mean accuracy scores in Table 2 show, there was a decrease in accuracy as the length of the string increased. The linear trend component for this main effect was significant, $F(1, 456) = 199.5$, and accounted for 97% of the sums of squares due to this factor. Thus, overall, as the number of letters in the string increased, accuracy decreased.

Significant first-order interactions of Nationality × Letter Length, $F(6, 456) = 14.68$, and Grade × Letter Length, $F(6, 456) = 4.25$, and the significant second-order interaction, $F(6, 456) = 5.23$, indicated that accuracy decreased differentially in the four groups of subjects. The difference in the linear trend component of the Nationality × Letter Length interaction was significant, $F(1, 456) = 82.32$, and accounted for 93% of the sums of squares due to the interaction: As letter length increased, accuracy showed a sharper decline for the American than for

[5] In evaluating the significance of the F values obtained in the simple effects tests, Dunn's method was used to control for the familywise Type I error probability. The interactions involved four means. Four pairwise comparisons represent the simple main effects. The use of Dunn's procedure (Keppel, 1973) and an alpha level of .05 familywise Type I error probability with four pairwise comparison requires a p value less than .0125 for significance.

Table 2
Mean Percentage of Accurately Pronounced Pseudowords and Correlation–Regression Analyses

Measure	Letter length 3	4	5	6	7	8	9
% correct	96.15	91.21	87.96	84.59	79.49	73.49	74.33

Correlation–regression analyses

Group	r	Slope	Mean % correct
American first graders	−.96*	−8.79	59.24
American third graders	−.96*	−4.04	87.22
Turkish first graders	−.82*	−1.31	93.64
Turkish third graders	−.92*	−1.50	94.45

* $p < .01$.

the Turkish subjects. The difference in the linear trend component was significant for the Grade × Letter Length interaction, $F(1, 456) = 17.26$, accounting for 68% of the sums of squares due to this interaction: Accuracy showed a sharper decrease for the first graders than for the third graders. The difference in the linear trend components was responsible for 62% of the sums of squares due to the three-way interaction of Nationality × Grade × Letter Length and was significant, $F(1, 456) = 19.52$. The nature of the linear trends for each group is shown in Figure 2.

The linear trend within each group was significant, all $Fs(1, 456) > 5.57$, $ps < .05$. Correlation-regression analyses were conducted on the group mean data, and the least squares regression lines for each group are indicated in Figure 2. Each of the correlations was significant. In order to test for differences in the rate at which accuracy decreases, the slopes shown in Table 2 were tested against one another. The decrease in accuracy was steeper for the American first graders than for the Turkish first graders, $t(36) = -5.5$. The American first graders also showed a sharper decrease than the American third graders, $t(36) = -5.28$. Slope comparisons among the other groups were nonsignificant (all $ps \geq .05$).

Thus, the latency data that indicated a lower cost per additional letter for the Turkish subjects were not due to a sharper decrease in accuracy among the Turkish subjects as compared with the American subjects. In fact, as the data in Figure 2 indicate, accuracy levels for the Turkish first graders were above 88%, and for the third graders, above 90%, whereas accuracy for the American first graders dropped to 38% and for the American third graders to 75% at the longest letter lengths. The American subjects tended to be both slower and less accurate than the Turkish as letter length increased. This was especially true for the first graders.

The ANOVA of the accuracy data also indicated that third graders were more accurate (91.3%) than first (76.4%), $F(1, 76) = 29.57$. Thus, as hypothesized, decoding accuracy increased across grades. There was, however, a significant Nationality × Grade interaction, $F(1, 76) = 22.81$. The means for this interaction are shown in Table 2. Simple main effects tests indicated that accuracy increased significantly only for the American subjects: Third graders were more accurate than first, $F(1, 76) = 7.45$. Accuracy differences in the Turkish subjects were nonsignificant. It was also the case that the main effect of nationality was due only to the first-grade subjects: Turkish first graders were more accurate than American, $F(1, 76) = 11.26$. Third-grade Turkish and American subjects were equally accurate.

The accuracy data thus indicate that the greater regularity of the Turkish language leads to high levels of correct decoding by the

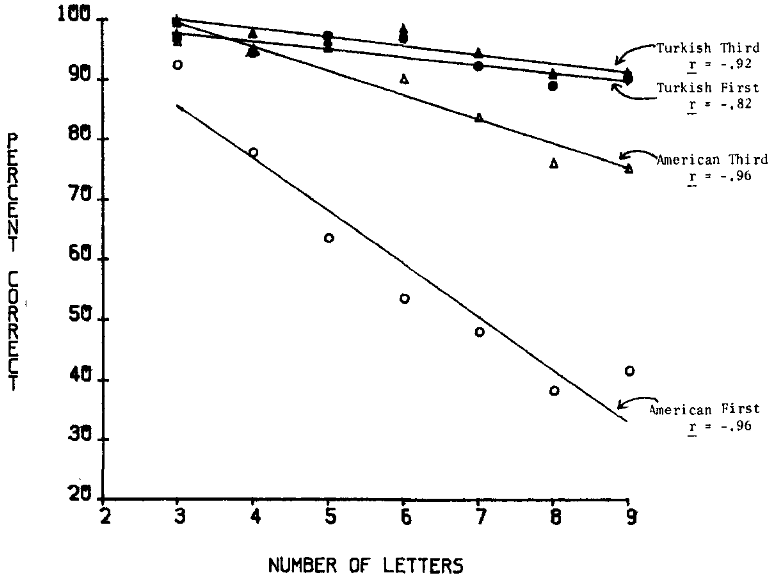

Figure 2. Relation between percentage of correct pronunciation and the number of letters in the pseudowords for each grade and nationality group.

end of first grade. Changes from first to third grade reside in increased speed of decoding. For the American students, both accuracy and speed of decoding increase from first to third grades. It is also interesting that although accuracy levels are equivalent in American and Turkish third graders, there exist large and significant differences in decoding speed, even when only correctly pronounced strings are considered. Similar results in decoding accuracy were reported by Venezky (1972) for students learning to read in Finnish, another language with high letter–sound correspondence. Kyostio (1980) also reported regularity of letter–sound correspondences to be a facilitating factor in reading acquisition, at least as far as reading skills such as decoding are concerned. Thus, the regularity in letter-sound correspondence in the Turkish language seems to facilitate the acquisition of decoding skills. This is reflected in both speed and accuracy of decoding.

The American subjects also had four additional vocalization latencies as a result of the nonoverlapping syllable types discussed in the Method section. These were submitted to an ANOVA that included grade (first and third) as the between-subjects factor and 18 syllable types as a within-subjects factor. Syllable type was significant, $F(17, 623) = 31.48$. Contrasts on the means indicated that for equivalent letter length, the nonoverlapping types were not significantly different from the overlapping types, $0.08 < Fs(1, 623) < 3.44$, all $ps > .05$. The means for each comparison, given in seconds, were as follows: one-syllable, three-letter overlap (1.86) versus nonoverlap patterns (1.81); one-syllable, four-letter overlap (2.52) versus nonoverlap patterns (2.43); five-letter, two-syllable, overlap (3.04) versus one-syllable, nonoverlap patterns (2.73); two-syllable, four-letter overlap (2.66) versus nonoverlap patterns (2.43).

These results indicate that the number of letters in the stimulus array is more important than the nature of the letter combinations used. Initially, we anticipated that the nonoverlapping types might behave differently than the overlapping types because the nonoverlapping patterns strongly reflect sources of irregularity in letter–sound correspondence in English. It might have been the case that these nonoverlap patterns would have shown decoding difficulty, whereas the patterns that overlap with Turkish would not. The finding that the overlap and nonoverlap patterns behaved similarly indicates that although irregular letter–sound correspondence may be restricted to only some patterns in a language, the effect is to increase the difficulty of acquiring decoding accuracy and speed for all patterns in the language.

Comprehension

The mean number of correct responses on the comprehension tasks was also computed for each subject. The maximum score was five on each text. A 2 (nationality) × 2 (grade) × 2 (text level) ANOVA with repeated measures on text level was conducted. Table 3 displays the results for the two texts at each grade and nationality level.

The analysis of the comprehension scores indicated that Turkish students scored higher ($M = 2.96$) than American students ($M = 2.40$), $F(1, 76) = 11.76$. Third graders scored higher ($M = 3.56$) than first graders ($M = 1.80$), $F(1, 76) = 115.41$. There was also a significant interaction of Nationality × Grade, $F(1, 76) = 29.26$. Simple main effects tests were conducted on this interaction. In the American sample, third graders scored higher ($M = 3.73$) than first graders ($M = 1.08$), $F(1, 76) = 65.26$. The performance of third graders in the Turkish sample ($M = 3.40$) was also significantly better than that of the first-grade Turkish students ($M = 2.53$), $F(1, 76) = 7.04$. Thus, as we had expected, comprehension scores increased over grades for both Turkish and American subjects. The Nationality × Grade interaction was due to between-nationality comparisons within grades: Turkish first graders ($M = 2.53$) scored higher than American first graders ($M = 1.08$), $F(1, 76) = 19.54$, but the performance of American ($M = 3.73$) and Turkish third graders was equivalent, $F(1, 76) = 1.01$, $p < .05$.

There was also a significant effect of text level, $F(1, 76) = 20.83$: The mean correct score on the primary-level text (2.94) was greater than that on the elementary-level text (2.43). Text level also interacted with grade, $F(1, 76) = 4.47$, $p < .05$. Simple main effects tests on this interaction indicated that text level was a significant factor only for the first graders. Their performance on the primary-level texts ($M = 2.18$) was higher than on the elementary-level texts ($M = 1.43$), $F(1, 76) = 14.24$, whereas the third graders performed equally well on the two text levels ($M = 3.70$ for primary and 3.28 for elementary), $F(1, 76) = 1.85$, $p < .05$.

Table 3
Mean Number of Correct Responses on Comprehension Task

	Grade			
	American		Turkish	
Text	1	3	1	3
Primary level	1.55	3.95	2.80	3.45
Elementary level	0.60	3.50	2.25	3.35

Thus, our hypotheses about comprehension were confirmed in that comprehension performance increased over grades. Turkish subjects at the first-grade level performed better on the comprehension task than did American first graders. This advantage was no longer present at the third-grade level, where the influence of decoding proficiency on comprehension was expected to be smaller.

Relationship Between Decoding and Paragraph Comprehension

In order to study directly the relationship between decoding speed and comprehension, correlations were computed between each subject's mean vocalization latency for correctly pronounced pseudowords and number of correct responses on the comprehension task. For both Turkish first and third graders, the correlation was significantly greater than zero ($rs = -.43$ and $-.53$, respectively, $ps < .05$, $df = 18$). For American first graders, the correlation was marginally significant ($r = -.344$, $p < .10$). The correlation for American third graders ($r = -.003$) was not significantly different from zero. Consistent with previous researchers (e.g., Curtis, 1980), we had expected to find a decreased relationship between vocalization latency and comprehension in the older American students. However, we were somewhat puzzled by the zero correlation in the American third graders because Curtis (1980) did find a significant correlation between pseudoword vocalization speed and reading comprehension grade level in Americans of this age level.

We examined whether the performance

linkage between decoding and comprehension was reflected in the vocalization errors. Only in the Turkish first graders was the correlation between vocalization accuracy and number correct on the comprehension measure significantly different from zero. The correlation coefficients were as follows: Turkish first graders, $r = .394$; Turkish third graders, $r = .147$; American first graders, $r = .258$; and American third graders, $r = .09$. (The critical value of r with 18 degrees of freedom = .378, $p < .05$.)

Thus, neither the vocalization latency nor the vocalization accuracy performance was correlated with comprehension in the American third graders. However, the comprehension measure used in the present study may not have been as sensitive as the measure used by Curtis (1980).[6] What is significant, however, is that, despite a potential lack of sensitivity in this measure, for the Turkish students there were significant correlations between vocalization latency and comprehension performance. These data would seem to indicate that for languages with high degrees of letter–sound correspondence, greater automaticity in decoding is associated with better comprehension. In languages where letter-sound correspondence is less regular, there may be a weaker association between decoding and comprehension.

The difference between the language groups in the degree of association between decoding proficiency and comprehension may implicate differences in the strategies for getting meaning adopted by the two groups. When there are high degrees of regularity in letter-sound correspondences, decoding proceeds quickly at high levels of accuracy. Individual word decoding thus appears to be an efficient route to meaning for languages such as Turkish. For English, however, individual word decoding not only appears to take more time but accuracy levels are lower, especially during the beginning phases of acquisition. Decoding as a route to meaning appears to be less efficient. In the face of less proficient letter-by-letter decoding skills, the American students may develop other strategies for determining meaning. One such strategy might be a heavier reliance on context cues for individual word decoding. Although the Turkish and American third graders did not differ on paragraph comprehension, the former group's decoding proficiency was superior to the latter's. It may well be that the American students' comprehension reflects a reliance on different strategies for deriving meaning from text. Having to use such a strategy relatively early in the acquisition of reading might produce benefits in comprehension when materials to be read include many unfamiliar words or vocabulary that is new to the students. Turkish students, or students reading other languages with regular letter-sound relationships, might have greater difficulty understanding such material. Research in the area of context cues to meaning is needed to further elucidate the consequences of learning to read a language with high versus low degrees of letter-sound correspondence.

We recognize that cross-cultural research is often fraught with many difficulties, most notably that of sample comparability. In the present research, every effort was made to find samples that were comparable and differed only with respect to language and culture. Despite such difficulties, cross-cultural research does provide an opportunity to examine questions of the types discussed in the present research in ongoing educational contexts. This is in contrast to research that employs artificial alphabets and examines learning over relatively short periods of time in the laboratory setting. The optimal situation exists when the findings from research done in a cross-cultural setting, such as the current work, converge with those obtained under the more controlled laboratory, single-culture settings.

---

[6] There are also several other differences between this study and that conducted by Curtis (1980). In the Curtis study, pseudoword vocalization was measured only on CVC and CVCC patterns, which constituted only two of the one-syllable, three- and four-letter patterns that were examined in the present study. Also note that although Curtis found a significant simple correlation between pseudoword vocalization speed and reading grade level in her third-grade sample ($r = .44$, $df = 38$, $p < .01$), multiple regression analyses indicated that for the third graders the three variables that led to signifiant improvement in accounting for the amount of variance in reading grade level did not include the pseudoword vocalization speed measure.

References

Chall, J. S. (1979). The great debate: Ten years later, with a modest proposal for reading stages. In L. B. Resnick & P. Weaver (Eds.), *Theory and practice of early reading* (Vol. 1, pp. 29–55). Hillsdale, NJ: Erlbaum.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72*, 656–669.

Frederiksen, J. R. (1978). Assessment of perceptual, decoding, and lexical skills and their relation to reading proficiency. In A. M. Lesgold, J. W. Pellegrino, S. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction* (pp. 153–169). New York: Plenum Press.

Frederiksen, J. R. (1980). Component skills in reading: Measures of individual differences through chronometric analysis. In R. E. Snow, P. A. Federico, & W. E. Montague (Eds.), *Aptitude, learning and instruction* (Vol. 1, pp. 105–138). Hillsdale, NJ: Erlbaum.

Keppel, G. (1973). *Design and analysis: A researcher's handbook.* Englewood Cliffs, NJ: Prentice-Hall.

Kyostio, O. K. (1980). Is learning to read easy in a language in which the grapheme-phoneme correspondences are regular? In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading and dyslexia* (pp. 35–49). Baltimore, MD: University Park Press.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.

Lesgold, A. M., & Perfetti, C. A. (1981). Interactive processes in reading: Where do we stand? In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 387–407). Hillsdale, NJ: Erlbaum.

Liberman, I. Y., & Shankweiler, D. (1979). Speech, the alphabet and teaching to read. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 2, pp. 109–132). Hillsdale, NJ: Erlbaum.

Malmquist, E. J., & Grundin, H. U. (1980). Cross-national studies on primary reading: A suggested program. In J. F. Kavanaugh & R. L. Venezky (Eds.), *Orthography, reading and dyslexia* (pp. 121–133). Baltimore, MD: University Park Press.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (1977). Language comprehension and fast decoding: Some psycholinguistic prerequisites for skilled reading comprehension. In J. T. Guthrie (Ed.), *Cognition, curriculum, and comprehension* (pp. 20–41). Newark, DE: International Reading Association.

Perfetti, C. A., & Lesgold, A. M. (1979). Coding and comprehension in skilled reading and implications for reading instruction. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 1, pp. 57–84). Hillsdale, NJ: Erlbaum.

Rozin, P., & Gleitman, L. R. (1977). The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principle. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading* (pp. 55–141). Hillsdale, NJ: Erlbaum.

Stanovich, K. E. (1981). Relationships between word decoding speed, general name-retrieval ability, and reading progress in first-grade children. *Journal of Educational Psychology, 6*, 809–815.

Venezky, R. L. (1967). English orthography: Its graphic structure and its relation to sound. *Reading Research Quarterly, 2*, 75–106.

Venezky, R. L. (1970). *The structure of English orthography.* Paris: The Hague.

Venezky, R. L. (1972). *The letter-sound generalizations of first, second, and third grade Finnish children* (Tech. Rep. No. 219). Madison: Wisconsin Research and Development Center for Cognitive Learning.

(*Appendix follows on next page*)

Appendix A

Pseudowords Used in the Vocalization Task

One-syllable pseudowords		
nat	sar	tig
pel	ber	tol
est	olt	ork
spom	gral	trab
grav	frez	draf
mank	mest	fulp
vank	garf	mers
Two-syllable pseudowords		
taba	reto	noro
devi	tada	mota
nurid	gemus	gomal
velin	mavem	danal
atlic	enher	atmas
alson	evken	atgat
permal	zurmen	sadbar
tervez	hersan	remsel
pingrom	destrul	fostvan
furpniz	nelgraf	mersfiz
steplon	trinmet	grapsel
prentip	trakpor	stadlum
Three-syllable pseudowords		
orliyal	anruver	ilvetal
orsaman	elkiyel	erdemet
saderval	tenemser	ruferlan
davansar	sevimrel	ravarnik
ripsolat	serbigal	vemrotal
levbetim	serviyal	banpotan
surdenper	lertonpel	tasverpul
mevtimler	tarvonsan	baltansar
Nonoverlap pseudowords		
fea	mie	boe
nea	ree	loe
tead	zeer	cail
poun	taul	noot
lave	vate	fube
tine	gase	dote
thros	strim	plib
scrat	spram	shren

Appendix B

Texts and Questions Used to Assess Comprehension

Primary-Level Text

Thousands of years ago, builders discovered clay dirt for making bricks. Builders used bare feet or a hoe to mix the clay dirt with straw and water. The mixture was poured into boxes and dried in the sun to become bricks. The bricks were then taken from the boxes and stuck together with fresh mud. If clay bricks were built on a strong base, they could stand for hundreds of years. Sun baked brick houses cost little money to build. They kept out heat in the daytime and cold at night. Although they are now made by machine, bricks continue to be important in the building business.

Questions:

1. When did builders discover clay dirt for making bricks?
2. What did the builders use to mix the clay dirt with straw and water?
3. How was the clay and water mixture dried?
4. What characteristics did the houses built with the sun baked bricks have?
5. What are bricks made by today?

Elementary-Level Text

For centuries, the circus has brought enjoyment to countless millions throughout the world. However, many of the features of this modern amusement were not always events in a "circus." Originally, the term was used by ancient Romans to describe the ring in which public games and important events were performed. Our modern circus originated in England about 200 years ago. Originally, it had only one ring where the audience would watch horseback riders perform tricks. Later, colorful costumes, amusing clowns, trained animals, and trapeze artists were introduced. The additional events made it necessary to increase the number of rings. However, only the middle ring, where all activities were performed, was considered the "circus." Today the circus is a general term describing all the people, amusements and events.

Questions:

1. Who used the term *circus* first?
2. Where did our modern circus originate?
3. What kind of amusements did the audience watch in the circus?
4. When additional events were introduced, what had to be done?
5. What does the term *circus* mean today?

Received March 14, 1983
Revision received October 3, 1983 ■