

REVIEW OF APPLIED LINGUISTICS

Published by the Afdeling Toegepaste Linguïstiek at the Katholieke Universiteit Leuven (Belgium), with the financial assistance from the Vlaamse Deelregering, the Universitaire Stichting and the Faculteit Letteren en Wijsbegeerte of the K.U. Leuven.

EDITOR IN CHIEF : Nicole Delbecq.

EDITORIAL COMMITTEE : R. De Bleser, M. Debrock, J. De Kock, L.K. Engels, D. Geens, M. Goethals, W. Martin and K. Van den Eynde.

MANAGING EDITOR : Fred Dekoning, ITI, Blijde-Inkomststraat 21, B-3000 Leuven.

EDITORIAL ADVISORY BOARD : Hugo Brandt Corstius, Universiteit van Amsterdam and Erasmus Universiteit Rotterdam, the Netherlands; John B. Carroll, Psychometric Laboratory, University of North Carolina, Chapel Hill, North Carolina, U.S.A.; Michael G. Clyne, Modern Languages Department, Monash University, Victoria, Australia; René Collier, U.F.S.I.A., Universiteit Antwerpen, Antwerpen, Belgium; Robert M. De Keyser, Department of Linguistics, University of Pittsburgh, Pittsburgh, U.S.A.; Joshua A. Fishman, Yeshiva University, New York, U.S.A.; Stephen Krashen, Linguistics Program, University of Southern California, Los Angeles, California, U.S.A.; Wolfgang Kuhlwein, Anglistik : Linguistik, Universität Trier, Trier, West Germany; Gerhard Nickel, Institut für Linguistik, Universität Stuttgart, Stuttgart, West Germany; Burghard Rieger, Institut für Mathematisch-Empirische Systemforschung, Aachen, West Germany; William C. Ritchie, Faculty of Foreign Languages, Syracuse University, Syracuse, New York, U.S.A.; Maurice van Overbeke, Institut de Linguistique, Université Catholique de Louvain, Louvain-

Is Spelling Acquired or Learned ? A Re-analysis of Rice (1897) and Cornman (1902)

Stephen D. KRASHEN
University of Southern California

Howard WHITE
Saniku Gakuin College

In this paper, we present a re-analysis, or "secondary analysis" of research results originally published over 80 years ago. A second look at this data is justified because the issues that it addresses are still far from settled: Is spelling learned implicitly or explicitly? (In other terms: Is spelling "caught" or "taught"? (Peters, 1985), or is spelling "acquired" or "learned"? (Krashen, 1982)).

The two research reports we will re-analyze provide a good opportunity for shedding some additional light on this question. The authors addressed a central issue, and, in our view, attempted to deal with this issue appropriately. As we shall see, subsequent studies have not improved on their methods. Also, in both cases, the authors presented a considerable amount of their data, data that to our knowledge has never been submitted to statistical analysis.

In our re-analysis, we applied common statistical tests to both Rice's and Cornman's data. Our analysis, for the most part, confirms both Rice's and Cornman's claims that formal instruction in spelling has limited effects. There are, however, some surprises,

some instances in which statistical analysis reveals somewhat different results from those claimed by Rice and Cornman. In the final section of this report, we attempt to account for the results of our re-analysis in terms of two central hypotheses, claiming that both implicit and explicit learning of spelling take place, but that the former is much more powerful. We also attempt to relate our hypotheses to more current studies of spelling.

Re-Analysis of Rice (1897)

Rice's study has been acknowledged as the first empirical investigation of the efficacy of spelling instruction (see e.g. Tidyman, 1915; Thompson, 1930). Rice concluded, on the basis of visual inspection of his data, that no relationship existed between the amount of time devoted to spelling instruction in the elementary school and spelling achievement. While this conclusion was accepted even by Rice's critics (Tidyman, 1915), it has never, to our knowledge, been verified using statistical techniques.

Rice used three tests, a "Column Test," a "Sentence Test," and a Composition.

The Column Test

The Column Test was simply a list of fifty words (e.g. "furniture," "chandelier," "independent") sent to "school superintendents in various sections of the United States" (p. 164). Twenty schools agreed to administer the test, and the identical list was used in grades four through eight with a total of about 16,000 children. Tests were administered by classroom teachers, who dictated the words in isolation (i.e. no sentence context) to the children.

Column Test results produced what Rice considered to be unacceptably high amounts of variation:

"On directing our attention to the results of the first test we are startled by the enormous variations, when the extremes are considered, - particularly in the fourth-year classes, where the averages range from 33 to 95.3 per cent ..." (pp. 164-165).

A two-month tour of some of the schools and visits with the teachers revealed what Rice concluded was the cause of this variability: Some teachers, apparently, gave students hints as to the correct spelling of certain words by exaggerating their pronunciation. While Rice considered the results of the Column Test to be invalid, he published them anyway, explaining that "... to ignore this test might have submitted me to the charge of hiding results which did not substantiate my pedagogical theories; whereas, in truth, I had fully determined, in advance, to base my theories on the facts, whatever their nature might be" (p. 186).

Rice's Table 2, from part I of his paper, presents average (median) scores for each school tested, along with the number of minutes per day devoted to spelling instruction in each school. Our Table 1 contains our calculations of means and standard deviations of Column Test results, based on Rice's Table 2.¹

Table 1: Column Test Results: Rice, 1897

grade	n	mean	standard deviation
4	15	57.86	17.40
5	14	70.65	9.27
6	17	79.95	7.83
7	17	83.67	5.26
8	12	86.52	5.60

mean = mean percent correct
n = number of schools

Inspection of the standard deviation in grade four in Table 1 shows the variability of scores that Rice was concerned about. This variation, however, is much less after grade 4.

In Rice's Table 2, the amount of time devoted to spelling instruction is reported for each school. In our Table 2, we present correlations between median scores provided by Rice for individual schools on the Column Test and the amount of time each school devoted to spelling instruction. Two kinds of correlations are presented. First, we simply correlated, for each grade level, average (median) scores provided by Rice for individual schools on the Column Test with the amount of time devoted to spelling instruction at the grade level ("simple correlations"). Second, we correlated Column Test scores with the cumulative time devoted to spelling instruction ("cumulative correlations"). For example, school 1C in Rice's Table 2 reported 45 minutes per day for spelling instruction in grade 4 and 35 minutes per day in grade 5. In calculating the simple correlation for grade 5, the 35 minute per day figure was used. To calculate the cumulative correlation for grade 5, we estimated the total time these children had been exposed to spelling instruction in grades 4 and 5. (Cumulative correlations are based on fewer schools than the simple correlations, due to missing data, and do not include data from grades 1, 2, and 3, since this information was not provided by Rice.)

Table 2: Correlations between Column Test scores and time devoted to spelling instruction (from Rice, 1897)

simple correlations			
grade	n	r	p
4	15	.432	ns (.11)
5	14	.791	<.01
6	17	.502	<.05
7	17	.496	<.05
8	12	.397	ns
cumulative correlations			
grade	n	r	p
5	12	.596	<.05
6	11	.811	<.01
7	10	.505	ns (.14)
8	10	.576	ns (.10)

The correlations in Table 2 are, surprisingly, all positive and, while not always statistically significant at the .05 level, are consistently high.² Despite Rice's conclusions, there does seem to be a relationship between the amount of time devoted to spelling and how well children do on the Column Test.

The Sentence Test

Rice reports that he personally supervised the administration of the Sentence Test. In this test, words were read by examiners in sentences, e.g. for the target words 'running', 'slipped', 'listened', 'queer', 'speech' and 'believe', students would hear:

While running, he slipped. I listened to his queer speech, but I did not believe any of it.

Grades 4 and 5 received a 50 word version of the Sentence Test, while grades 6, 7, and 8 did a 75 word test that contained most of the grade 4 and 5 words as a subpart. Rice noted that "in preparing

this test, special care was exercised to omit words whose pronunciation would tell the secret" (p. 166).

Table 3 presents means and standard deviations for Sentence Test scores. According to Rice, over 13,000 children took this test. (The fact that fifth graders do better than sixth graders in Table 3 is probably due to the differences in tests used.)

Table 3: Sentence Test Results: Rice, 1897

grade	n	mean	standard deviation
4	16	68.50	5.40
5	15	78.13	2.95
6	16	73.31	4.33
7	14	80.36	4.58
8	12	86.67	3.60

n = number of schools

All scores in terms of percent correct; for grades 4 and 5, perfect score = 50; for grades 6, 7, and 8, perfect score = 75.

Table 4 presents simple and cumulative correlations between Sentence Test results and the amount of time devoted to spelling.

Table 4: Correlations between Sentence Test results and time devoted to spelling instruction: Rice, 1897

grade	n	r	p
4	16	-.063	ns
5	15	-.266	ns
6	16	.287	ns
7	14	.104	ns
8	12	-.065	ns

grade	n	r	p
5	13	-.201	ns
6	12	.290	ns
7	9	-.240	ns
8	8	-.305	ns

This time, correlations are low and are mostly negative, without a hint of statistical significance, in agreement with Rice's characterization of his data and in contrast to the Column Test results. We will discuss possible reasons for this discrepancy later.

The Composition Test

In this test, children were read a story, which was accompanied by a picture, and were asked to reproduce the story. Rice does not mention what the story was about, or whether the same story was used with all grade levels. He does, however, note that the compositions "... in a large number of schools, were prepared under my personal supervision, and in many others under the special supervision of persons whose conscientiousness could not be doubted" (p. 187).

Results of the Composition Test are presented in terms of the percentage of words spelled correctly. Rice points out that when the same word was misspelled more than once in a composition, only one error was counted (p. 167); as we shall see below, Cornman (1902) used a slightly different method. Rice did not include data on the total number of words produced.

Our calculations of mean scores and standard deviations for the Composition Test are presented in Table 5.

Table 5: Composition Test results: Rice, 1897

grade	n	mean	standard deviation
4	10	97.49	.599
5	10	98.31	.418
6	8	98.45	.444
7	7	98.93	.265
8	9	99.14	.368

Scores on the Composition Test are remarkably high; students at all grades can, apparently, spell correctly most of the words they generate themselves in their writing, even though the topic was not self-generated. Despite the very high scores, there is a clear progression from grade 4 to grade 8, with higher grades doing better; in fact, the rank order correlation between accuracy and grade level is perfect.

Table 6: Correlations between Composition Test scores and time devoted to spelling instruction

simple correlations			
grade	n	r	p
4	10	-.291	ns
5	10	-.788	<.01
6	8	.638	ns (.09)
7	7	.252	ns
8	9	-.319	ns
cumulative correlations			
grade	n	r	p
5	9	-.613	ns (.08)
6	7	.444	ns
7	6	-.089	ns
8	8	.171	ns

Correlations between Composition Test results and time devoted to spelling instruction (Table 6) show no clear pattern. Some are positive, some negative, and some are small, with only one reaching statistical significance.³

Summary of Rice (1897)

The following table summarizes the results of our analysis of the three tests used by Rice:

test	relationship to amount of instruction
Column Test	positive
Sentence Test	none
Composition	none

With the exception of the Column Test results, our re-analysis agrees with Rice's characterization of his data.

Based on the results of the Sentence Test and the Composition Test, Rice interpreted his results as showing "the futility of the spelling grind," and recommended that no more than 15 minutes daily be devoted to spelling instruction (p. 416). We will re-evaluate these conclusions after examining a second study of spelling, Cornman's 1902 monograph, *Spelling in the Elementary School*.

Re-Analysis of Cornman (1902)

Oliver Cornman, the principal of the Northwest School in Philadelphia, supervised an investigation designed to determine to what extent spelling proficiency was related to spelling instruction. We will first discuss the portion of Cornman's research that parallels Rice's experiment and then move to Cornman's major focus: In

what would be considered today a bold move, Cornman suspended all spelling instruction and all spelling homework in his school for three years and examined the effect of student spelling proficiency.

PART I: CORNMAN'S REPLICATION OF RICE

Cornman, in his Table 19, provides data on the relationship between spelling proficiency as measured by the Term Examination, and the amount of time devoted to spelling instruction in his school (Northwest), another school in which spelling instruction was dropped (the Agnew school), and eleven other schools in the Philadelphia area.⁴ Cornman gives little information about the Term Examination, other than the fact that it is "the regular term examination in spelling (which is prepared in the office of the superintendent and is the same for all the schools of Philadelphia) ..." (p. 59).

Cornman claimed that the data in his Table 19 show no relationship between the amount of time devoted to spelling and spelling proficiency. According to our calculations, however, this was not entirely the case. As our Table 7 shows, the simple correlation between amount of spelling instruction and spelling proficiency is fairly substantial for grade 3, and the cumulative correlations, while not reaching statistical significance, are consistently positive and are quite high.⁵

Table 7 Correlations between time devoted to spelling instruction and spelling proficiency (Term Examination): Cornman, 1902

simple correlations			
grade	n	r	p
3	10	.435	ns
4	10	.145	ns
5	9	.147	ns
6	11	.127	ns
7	11	-.165	ns
cumulative correlations			
grade	n	r	p
4	8	.531	ns (.17)
5	7	.434	ns
6	8	.538	ns (.17)
7	7	.529	ns (.22)

n = number of schools

Looking at the Rice and Cornman data together, we are left with this situation: Consistently positive correlations between the amount of spelling instruction and test performance for Rice's Column Test, positive cumulative correlations for Cornman's Term Examination, and no correlation between instruction and spelling for Rice's Sentence Test and Composition Test.

PART II: THE EFFECT OF DROPPING SPELLING INSTRUCTION

We turn now to the second part of Cornman's study, the effect of dropping formal spelling instruction from the Northwest and Agnew schools. The treatment lasted for three years:

September, 1897, to June, 1898
September, 1898, to June, 1899
September, 1899, to June, 1900

During this time, students received no spelling lessons and had no spelling homework. Cornman does note, however, that spelling errors in written work were corrected, and "pupils were taught to appeal, when in doubt, to the teacher and ... to the dictionary" (p. 67).

Cornman used a variety of measures to see whether the lack of spelling instruction made a difference: the "15 Minute Test," several used by Rice (Column Test, Sentence Test, and Composition), and the Term Examination.

15 Minute Test

In the 15 Minute Test, students were simply asked to write lists of words for 15 minutes: "To obtain the lists of spontaneously selected words, the pupils were given conveniently prepared slips of paper and told to begin at a given signal and write as many words as they could, using any words at all that 'they might happen to think of.' The words were written in vertical columns, and at the end of fifteen minutes the signal to stop writing was given and the papers were collected" (p. 7).

Unfortunately, 15 Minute Tests were not given for the entire duration of the project, but were only given one year before the

treatment began (June, 1896), just before it began (June, 1897), and one year later (June, 1898). We thus have two comparison groups who had traditional spelling instruction (tested June, 1896; June, 1897) and one experimental group (tested June, 1898), the experimental group having had one year of no spelling instruction.

In his Tables 3, 5, and 7, Cornman presents average (median) percent correct scores for boys and girls in grades 3b, 3a, 4b, 4a, 5b, 5a, 6, and 7 for the 15 Minute Test (recall that 3b refers to the first semester of third grade, 3a to the second semester). Since Cornman's monograph is not widely available, our re-arrangement of his data is presented in Table 8. We applied a One-Way Analysis of Variance for Repeated Measures to this data, and found no significant difference among the groups for boys and for girls (Table 9).⁶

Table 8 Northwest School 15 Minute Test Results

Boys:	grade	COMPARISON GROUP				EXPERIMENTAL GROUP			
		6/96		6/97		6/98			
		n	median	n	median	n	median	n	median
	3b	10	89.6	24	88.1	33	91.2		
	3a	24	92.6	19	91.4	21	91.6		
	4b	38	91.0	16	95.2	17	91.2		
	4a	14	94.5	23	90.1	21	91.7		
	5b	31	94.2	38	94.3	29	93.2		
	5a	31	92.2	24	92.3	32	92.7		
	6	40	96.2	53	92.6	55	93.8		
	7	29	97.0	26	96.2	31	94.4		
	8	34	98.1	33	95.8	31	97.3		

Girls: grade	COMPARISON GROUP		EXPERIMENTAL GROUP	
	6/96	6/97	6/98	
	n	n	n	median
3b	25	19	25	95.0
3a	19	20	17	94.6
4b	23	16	20	95.3
4a	18	23	22	95.9
5b	40	31	37	96.2
5a	43	18	32	96.1
6	11	69	34	97.1
7	36	28	36	96.7
8	36	46	36	98.4

n = number of students tested

All scores are percentages of words spelled correctly.

Table 9 Repeated measures analysis of variance:
15 Minute Test

Boys:	Sum of		Mean Square		F	p
	Squares	df				
Between groups	5.869	2	2.934		1.350	.287
Within grades	121.153	8	15.144			
Interaction	34.784	16	2.174			
Total	161.806	26				

Girls:

	Sum of	df	Mean Square	F	p
Squares					
Between groups	3.402	2	1.701	1.983	.169
Within grades	41.974	8	5.247		
Interaction	13.724	16	.858		
Total	59.1	26			

data from Cornman (1902), Tables 3,5,7

As indicated in Table 8, scores on the 15 Minute Test were very high. This suggests that a ceiling effect may be hiding real differences among the groups. While this is possible, the fact that accuracy in the 15 Minute Test clearly increases with age indicates that the scores have some validity. Another possible problem is the modest sample size; we will see, however, that similar results will be found for self-generated words in compositions with larger samples. In addition, it could be argued that self-generated words are not challenging to spell; writers may avoid those words that give them problems (Wallin, 1910). On the other hand, nearly all real-world spelling tasks involve self-generated words.

The Composition Test

Cornman provides extensive detail on composition scores, but gives us little detail on the conditions under which the compositions were written, and no information about topics. Length of compositions was modest, ranging from about 100 to 200 words. (The median length for grades 3 and 4 of the Northwest School on the June, 1897 test was 104 words. For grades 5 through 8, the median length was 189 words.)⁷

Cornman followed Rice's scoring procedure, with one exception: If a student misspelled the same word more than once, Cornman counted each instance as a separate error. As noted earlier, Rice only counted one error in these cases. The same composition, thus, might get a higher score using Rice's method, as compared to Cornman's.

Cornman provides data on compositions for June, 1897 through May, 1900. In Table 10, we present his data for tests administered at the Northwest School in June, 1897, 1898, and 1899, as well as results from May, 1900, and Rice's median composition data (from Cornman's Tables 15, 22, and 23). The June, 1897 and Rice data can be considered comparison groups; recall that spelling instruction was not dropped until September, 1897, and that the schools Rice studied all had spelling instruction, which ranged from ten to 50 minutes per day. The June, 1898, 1899, and May, 1900 groups are experimental groups, and show the effect of no spelling instruction on compositions after one, two, and three years respectively.

Table 10 The effect of dropping spelling instruction on Composition Test results

grade	COMPARISON GROUPS		EXPERIMENTAL GROUPS			
	Rice	med	6/97	6/98	6/99	5/00
3b			n med	n med	n med	n med
3			43 92.4	52 93.0	62 92.8	107 94.3
3a			33 96.3	36 96.2	48 95.5	
4b			30 97.5	46 94.0	34 96.1	
4						88 94.3
4a	97.5		48 95.0	43 98.6	39 97.0	
5b	97.8		73 97.3	70 97.5	57 97.2	
5						99 97.6
5a	98.4		42 97.6	63 96.2	55 98.6	
6b	98.1		65 97.7	56 97.5	56 98.1	
6						94 99.0
6a	98.4		46 97.9	29 98.5	44 98.9	
7	98.8		54 99.2	71 98.5	45 98.8	55 98.6
8	99.2		69 99.5	60 99.5	37 98.5	42 99.8

med = median

n = number of students tested

All scores are percentage of words spelled correctly.

It was possible to perform a One-Way Analysis of Variance for Repeated Measures on the June, 1897, June, 1898, and June, 1899 scores. No significant difference was found among these groups (Table 11). A second ANOVA was performed comparing June, 1897, June, 1898, June, 1899, and Rice's data. This was possible for grades 4a through 8 only. Again, differences were not significant.

Table 11 Repeated measures analysis of variance:
Northwest School compositions

June, 1897; June, 1898; June, 1899;

	Sum of Squares	df	Mean Square	F
Between groups	.269	2	.134	.136 ns
Within grades	94.994	9	10.555	
Interaction	17.831	18	.991	
Total	113.094			

June, 1897; June, 1898; June, 1899; plus data collected by
Rice (1897):

	Sum of Squares	df	Mean Square	F
Between groups	1.290	3	.430	.430 ns
Within grades	14.205	6	2.368	
Interaction	11.198	18	.622	
Total	26.198	27		

Data from May, 1900 could not be included in either analysis, since "a" and "b" scores for each year were combined. But visual inspection of Table 10 shows that these scores were similar to those of the other groups.

It thus appears to be the case that dropping spelling instruction for three years had no effect on composition scores.

Cohort analysis

It was possible to get a longitudinal picture of the progress of the students in Cornman's sample by rearranging data from his Tables

15 and 22, so that cohorts of students who did not get spelling instruction could be examined. This is done in our Table 12, which also includes the June, 1897 cross-sectional data for comparison.⁸ The data presented in this table is quite clear: Each cohort shows progress in spelling, Cohort 1 improves every semester, and the other cohorts show consistent (but not always perfect) progress. Each cohort also agrees closely with the comparison data.

Table 12 The effect of dropping spelling instruction
on Composition Test results: longitudinal data

comparison group	cohorts			
grade	6/87	1	2	3
3b	92.4	92.4 (6/97)		4
3a	96.3	95.7 (1/98)	96.3 (6/97)	
4b	97.5	96.2 (6/98)	97.3 (1/98)	97.5 (6/97)
4a	95.0	97.1 (1/99)	98.6 (6/98)	95.7 (1/98) 95.0 (6/97)
5b	97.3	97.2 (6/99)	97.9 (1/99)	97.5 (6/98) 97.2 (1/97)
5a	97.6		98.6 (6/99)	97.2 (1/99) 96.2 (6/98)
6b				97.0 (1/99)

Date of testing in parentheses.

All scores are average (median) percent correct.

Some potential problems

As noted earlier in our discussion of the 15 Minute Test, scores for compositions are high, which suggests that a ceiling effect might be operating. Such a ceiling effect might reduce or hide real differences among the groups. As was the case with the 15 Minute Test, however, there is a clear improvement from grades 3 to 8, which suggests that the finding of similarity among groups is genuine.

A second problem is fluctuating sample size. In Cohort 1, for example, 43 children took the Composition Test in grade 3b, 36 in grade 4b, and 57 in grade 5b. This could mean that there was high transiency at the Northwest School, or simply that not all of the children were able to take the test on a given day. If the latter was the case, and if random factors prevented some children from taking the Composition Test, Cornman's results are still valid. If the former is true, Cornman's conclusions that spelling instruction makes no difference still hold, but may not be valid for all three years, since a significant subset of the sample may not have been at Northwest the entire time.

Finally, as noted in our discussion of the 15 Minute test, compositions utilize self-generated words. As noted earlier, there may be differences between the results of tests utilizing self-generated words and tests requiring students to spell words given to them.

The Column Test

The Column Test used by Cornman was identical to that used by Rice. The same test was administered to grades 4a and up, with a slightly easier version used in grades 3 and 4b. Table 13, from Cornman's Table 20, presents Column Test results for boys and girls separately, with June, 1897 serving as the comparison group, and June, 1898 serving as the experimental group. Unfortunately, only the effect of one year of no spelling instruction could be measured.⁹ Application of t-tests for correlated samples showed no significant difference between instructed and non-instructed boys ($t = 1.162$, $df = 9$, ns) or between instructed and non-instructed girls ($t = .446$, $df = 9$, ns).

Table 13 Northwest School Column Test

Boys: grade	comparison group 6/97		experimental group 6/98	
	n	median	n	median
3b	31	45.1	32	40.5
3a	29	85.0	17	67.3
4b	13	79.1	19	77.0
4a	24	57.8	17	61.0
5b	39	68.9	38	64.7
5a	24	58.2	33	62.0
6b	30	74.6	37	71.3
6a	12	80.3	13	74.1
7	32	82.4	33	79.2
8	31	86.3	27	93.0

Girls:

grade	comparison group		experimental group	
	n	6/97 median	n	6/98 median
3b	19	55.2	20	43.5
3a	13	79.0	19	59.7
4b	17	81.0	27	88.4
4a	21	62.0	26	72.0
5b	34	67.4	37	69.8
5a	18	72.0	30	74.3
6b	29	77.7	19	78.0
6a	32	84.0	16	81.8
7	35	85.4	38	83.2
8	47	90.3	33	94.3

n = number of students tested

All scores are percentages of words spelled correctly.

Table 14 Column Test: boys and girls combined

grade	COMPARISON GROUPS		EXPERIMENTAL GROUP	
	Rice	6/97	6/98	
4a	59.6	59.9	66.3	
5b	65.5	68.2	67.3	
5a	69.0	65.6	68.2	
6b	75.0	76.2	74.7	
6a	80.0	82.2	78.0	
7	83.0	83.9	81.2	
8	87.6	88.3	93.7	

All scores are average (median) percentage of words spelled correctly.

Table 14 compares the combined Northwest data on the Column Test to Rice's results on the same test. The June, 1897 Cornman data and the Rice data are comparison groups, while the June, 1898 Cornman group is the experimental group.¹⁰ Repeated Measures Analysis of Variance showed no significant difference among these three groups (Table 15).

Table 15 Repeated measures analysis of variance:
Column Test

	Sum of Squares	df	Mean Square	F
Between groups	7.012	2	3.506	.604 ns
Within grades	1781.393	6	296.899	
Interaction	69.661	12	5.805	
Total	1858.066	20		

(see Table 14)

Cornman also collected data on the Column Test at the Agnew School (grades 3 and 4 only), presented in Table 16. There was no significant difference between comparison group boys and experimental boys ($t = .662$, $df = 3$, ns) and comparison and experimental girls ($t = 2.661$, $df = 3$, $p = .152$), but the t-ratio for the latter comparison is substantial, with comparison girls tending to outperform experimental girls (see footnote 2). When girls and boys are combined, comparison subjects tend to be better ($t = 3.484$, $df = 3$, $p = .078$).

Table 16 Agnew School Column Test

Boys:		Girls:	
grade	comparison 6/97	experimental 6/98	comparison 6/97
n	median	n	median
3b	20 52.4	22 51.7	32 50.7
3a	25 66.4	15 68.1	23 70.5
4b	13 73.0	16 64.6	21 75.8
4a	21 65.6	17 66.8	19 64.8

combined:

grade	comparison (6/97)	experimental (6/98)
n	median	n
3b	52 51.6	46 49.3
3a	48 68.9	37 66.3
4b	34 74.4	34 70.1
4a	40 65.2	27 64.4

n = number of students tested

All scores are percentage of words spelled correctly.

The Sentence Test

Here again, Cornman used the same test that Rice used. The upper grades (6, 7, and 8) were tested on 75 words, while the lower grades were tested on 50 words. In addition, the test given to grades 3 and 4b was slightly different from the version given to grades 4a and 5. Our Table 17 presents Sentence Test scores for Northwest boys and girls, extracted from Cornman's Table 20.

Statistical analysis shows no significant difference between comparison and experimental groups for girls (t-test for correlated samples; $t = .178$, $df = 9$, ns). For the boys, however, the

comparison group was superior ($t = 4.252$, $df = 9$, $p < .01$). For combined groups (Table 18), there was no significant difference.

Table 17 Northwest School Sentence Test

Boys:

grade	comparison group		experimental group	
	n	6/97 median	n	6/98 median
3b	24	56.9	32	54.3
3a	29	88.0	17	75.7
4b	13	84.2	19	76.4
4a	24	72.2	17	74.2
5b	39	79.9	38	76.4
5a	24	79.1	33	70.8
6b	30	70.9	37	66.8
6a	12	82.5	13	73.9
7	32	84.0	33	74.5
8	31	89.6	27	87.0

Girls:

grade	comparison group		experimental group	
	n	6/97 median	n	6/98 median
3b	19	58.7	20	61.0
3a	13	85.0	19	65.0
4b	17	87.3	27	91.1
4a	21	77.9	26	85.0
5b	34	77.7	37	81.9
5a	18	82.1	30	83.0
6b	29	74.5	19	76.5
6a	32	84.5	16	81.0
7	35	88.2	38	82.9
8	47	90.0	33	94.1

n = number of students tested

All scores are percentages of words spelled correctly.

Table 18 Northwest School Sentence Test:
boys and girls combined

grade	comparison group		experimental group	
	n	6/97 median	n	6/98 median
3b	43	57.8	52	57.7
3a	33	86.5	36	70.4
4b	30	85.8	46	85.3
4a	48	75.1	43	80.7
5b	73	78.8	70	79.2
5a	42	80.6	63	76.9
6b	59	72.7	56	71.7
6a	44	83.5	29	77.5
7	67	86.1	71	78.7
8	78	89.8	60	90.6

n = number of students tested

All scores are percentages of words spelled correctly.

The combined comparison and experimental groups were also compared to Rice's Sentence Test results for grades 4a to 8 in an Analysis of Variance for Repeated Measures. While the obtained F-ratio was significant ($F = 6.683$, $p = .011$), post-hoc comparisons showed that this was due to depressed performance by Rice's subjects, considered here to be a comparison group. Scheffe tests showed no significant difference between Cornman's experimental and comparison subjects ($t = 1.055$), agreeing with the results presented just above, but both Cornman groups outperformed Rice's subjects (experimental group versus Rice; $t = 2.504$, $p < .10$; critical value of $t = 2.64$ for the .05 level; comparison group versus Rice; $t = 3.56$, $p < .01$).

Sentence Tests were also administered at the Agnew School after one year of no spelling instruction (Table 19). Recall that the Agnew School only covers grades 3 and 4. Experimental boys were significantly better than comparison boys ($t = 5.238$, $df = 3$, $p < .05$), but there was no significant difference between experimental and comparison girls ($t = .647$). While the combined experimental group was not significantly better than the comparison group ($t = 2.465$, $df = 3$, $p = .18$), the combined experimental group was better at each level (see footnote 2).

Table 19 Sentence Test: Agnew School

Boys:

grade	comparison group		experimental group	
	n	6/97 median	n	6/98 median
3b	20	63.4	22	68.1
3a	25	68.8	15	72.2
4b	13	77.6	16	79.4
4a	21	74.7	17	79.5

Girls:

grade	comparison group		experimental group	
	n	6/97 median	n	6/98 median
3b	32	68.7	24	67.2
3a	23	75.8	22	75.2
4b	21	87.3	18	87.9
4a	19	78.8	10	84.4

Combined:

3b	52	66.1	46	67.7
3a	48	72.3	37	73.7
4b	34	82.5	31	83.7
4a	40	76.8	27	82.0

n = number of students tested

All scores are percentages of words spelled correctly.

The Term Examination

Three comparisons were possible using the Term Examination (Table 20). Recall that the Term Examination was a test given by the office of the Superintendent of Schools in Philadelphia and was given to all schools in the city. In his Table 26, Cornman compared the combined performance of students at the Northwest and Agnew schools to other schools in Philadelphia. In the first administration, given in January, 1898, the Northwest-Agnew group was compared to eleven other schools. By this time, the Northwest-Agnew group had been without spelling instruction for five months. Application of t-tests for correlated samples showed no difference between the groups ($t = 1.195$, $df = 4$). In June, 1899, the experimental group was compared to 19 comparison schools, after two

years of no spelling instruction. Statistical analysis again revealed no difference between the groups ($t = .156$, $df = 4$). In June, 1900, the Northwest and Agnew students, after three years of no spelling instruction, were compared to students at 50 Philadelphia schools. At this time, the comparison schools tended to do better on the Term Examination ($t = 2.728$, $df = 4$, $p = .053$, two-tailed test).

Table 20 Term Examination Results:
Northwest and Agnew Schools compared
to other Philadelphia schools

grade	January, 1898		June, 1899		June, 1900	
	comp.	exper.	comp.	exper.	comp.	exper.
3	67.6	62.8	69.6	65.2	75.0	69.0
4	63.3	70.5	72.8	68.7	76.0	67.8
5	71.4	72.3	66.8	80.5	71.6	72.6
6	73.7	76.6	81.4	80.0	70.4	65.1
7	75.8	82.9	69.1	68.0	73.0	69.9

number of schools in comparison groups:

1898: 11

1899: 19

1900: 50

Summary

Table 21 presents a summary of the Cornman data relating to the absence of spelling instruction, and, for convenience, repeats the findings on correlations between spelling proficiency and amount of spelling instruction.

Table 21 Summary of Results: Cornman (1902)

15 Minute Test (Table 8)

GRADES	COMPARISON GROUP	DURATION	RESULTS
boys	3-8 June, 1896, 1897	one year	no difference
girls	3-8 June, 1896, 1897	one year	no difference

Compositions (Table 10)

GRADES	COMPARISON GROUP	DURATION	RESULTS
3-8	June, 1897	two years	no difference
3-8	June, 1897, Rice	two years	no difference

Column Test (Tables 13,14,16)

GRADES	COMPARISON GROUP	DURATION	RESULTS
boys	3-8 June, 1897	one year	no difference
girls	3-8 June, 1897	one year	no difference
combined	4-8 June, 1897, Rice	one year	no difference

Agnew School

GRADES	COMPARISON GROUP	DURATION	RESULTS
boys	3&4 June, 1897	one year	no difference
girls	3&4 June, 1897	one year	no difference
combined	3&4 June, 1897	one year	comparisons tend to be superior (p = .078)

Sentence Test (Tables 17,18,19)

Northwest School

GRADES	COMPARISON GROUP	DURATION	RESULTS
boys	3-8 June, 1897	one year	comparisons superior
girls	3-8 June, 1897	one year	no difference
combined	3-8 June, 1897	one year	no difference

Agnew School

GRADES	COMPARISON GROUP	DURATION	RESULTS
boys	3&4 June, 1897	one year	exp. superior
girls	3&4 June, 1897	one year	no difference
combined	3&4 June, 1897	one year	no difference

Term Examination (Table 20)

(Northwest and Agnew combined; grades 3-7)
 COMPARISON GROUP DURATION RESULTS
 11 Philadelphia schools 5 months no difference
 19 Philadelphia schools two years no difference
 50 Philadelphia schools three yrs comparisons tend to be superior

Relation between spelling proficiency and
 amount of spelling instruction

Test	Results
Column Test (Rice)	positive correlation
Sentence Test (Rice)	no correlation
Compositions (Rice)	no correlation
Term Exam (Cornman)	positive cumulative correlations

Since the data are complex, we found it useful to resummariize the comparisons in Table 22. As indicated in Table 22, experimental-control comparisons favored comparison groups in only three cases. In two of these cases, the difference did not reach the .05 level, and in all cases, the differences between the groups were not large (the average comparison group advantage in the Column Test comparison was 2.6%, for the Sentence Test, 5.7%, and for the Term Examination, 3.9%). In one case, the experimental group was superior, and this difference was also small (3.7%). It should be also be noted that we performed over 20 experimental-comparison group comparisons; by chance, one out of 20 would be significant.

Table 22 Effects of dropping spelling

<u>No difference between groups</u>	<u>Comparison better</u>
15 Minute Test; 3-8, boys, 1yr	Column Test; 3&4, comb, 1yr
3-8, girls, 1yr	Sentence Test; 3-8, boys, 1yr
Compositions; 3-8, 1yr	Term Exam; 3-7, comb, 3yrs
3-8, 2yrs	
Column Test; 3-8, boys, 1yr	
3-8, girls, 1yr	
4-8, comb, 1yr	
3&4, boys, 1yr	
3&4, girls, 1yr	<u>Experimental better</u>
Sentence Test; 3-8, girls, 1yr	Sentence Test; 3&4, boys, 1yr
3-8, comb, 1yr	
3&4, girls, 1yr	
3&4, comb, 1yr	
Term Exam; 3-7, comb, 5m	
3-7, comb, 2yrs	

Discussion

In this section, we will attempt to account for Rice's and Cornman's results. Our explanation is based on the premise that both implicit learning ("acquisition") and explicit learning ("learning") of spelling occur. Below, we outline some hypotheses, attempt to show how they can account for the Rice and Cornman data, and argue that they are consistent with more recent studies of spelling.

Acquisition and Learning

It has been hypothesized (Krashen, 1981, 1982) that there are two distinct ways of developing ability in language. Language acquisition occurs subconsciously via comprehension of messages, and is represented subconsciously in the brain. It is similar to what has been termed "implicit learning." Language learning is "explicit learning"; it is the result of deliberate teaching and study and is represented consciously in the brain.

It has been argued that very little of language can be "learned" and learned competence plays a limited role in language production. We are able to use conscious learning only as a "Monitor," to edit the output of the acquired system. There are severe conditions on Monitor use: to use conscious rules to raise output accuracy, the performer needs to be focussed on form, needs to know the rule he is attempting to apply, and needs to have sufficient time to apply the rule (for evidence from second language acquisition and performance, see Krashen, 1981, 1982, 1985a).

The acquisition-learning framework can successfully account for a great deal of the Rice-Cornman data, as well as other data in spelling. Specifically, it hypothesizes that most spelling competence

is acquired by reading (despite Rice's findings; see footnote 3; for arguments against the hypothesis that spelling competence is the result of writing, see Smith, 1981, Krashen, 1990). It also makes the following hypotheses for performance on the spelling tests in Rice's and Cornman's studies:

1. Performance on compositions is the result of acquired knowledge. It can be argued that when students write compositions, they are not focussed on form. Rather, they are primarily concerned with meaning.
2. Performance on the 15 Minute Test is also the result of acquired competence. In the 15 Minute Test, students were, mostly likely, not focussed on form but were concentrating on writing down as many words as possible as rapidly as possible.

The Composition Test and the 15 Minute Test results are fully consistent with these hypotheses. We found no measurable impact of conscious learning on these tests: There was no correlation between test scores and time devoted to instruction for compositions (Rice) and no effect of dropping instruction for both measures (Cornman). We also hypothesize that for students in grade 3 and above, acquired competence is enough to allow them to write with good accuracy (although it may not be enough for first graders; see data from Clarke, 1988, in Table 23). This explains the high level of accuracy shown by Rice's and Cornman's subjects in compositions. Finally, we hypothesize that improvement shown in compositions and in the 15 Minute Test scores as children get older is a result of continued reading, not instruction. (We are hypothesizing that the spelling correction that Cornman's children received on their written work had little or no effect, a result consistent with more recent research on correction; for a review, see Krashen, 1990.)

3. Performance on the Sentence Test is also due to acquired competence. Words in the Sentence Test are presented in a

meaningful context; thus, students tended to focus on meaning rather than on form. In fact, students taking the Sentence Test, while aware they were taking a spelling test, may not have known which words were target words.

Consistent with hypothesis (3), Sentence Test performance does not correlate with amount of spelling instruction, and dropping spelling instruction has no consistent effect on Sentence Test scores.

4. Performance on the Column Test will show the effects of conscious learning in addition to the effects of acquisition. Column Test items are presented out of context, in lists, which focusses the test-taker on form.

Rice's data is consistent with hypothesis (4): There are clear positive correlations between amount of instruction and performance on the Column Test. In Cornman's data, however, it was found that dropping instruction generally had no impact on Column Test scores. One reason for this discrepancy may be the fact that the schools in Cornman's sample devoted less time to spelling instruction, averaging less than 15 minutes per day, while the schools Rice surveyed averaged about 30 minutes per day. Nevertheless, the correlations between Column Test scores and time devoted to spelling are strong enough so that 15 minutes per day should make a noticeable difference.¹¹ Perhaps, after nearly 90 years, we need some additional data.

5. As noted earlier, we have little information about the Term Examination, and the results, as was the case for the Column Test, appear to be in conflict: Positive correlations between Term Examination performance and time devoted to spelling, yet no effect for dropping instruction for five months and for two years. Instructed students did better after three years, however. It is possible that instruction has an effect, but it was too weak to be felt after five months or even two years (but see note 11).

While the data are not entirely consistent, we can conclude that where instruction has an effect, it is plausible that students were focussed on form (Column Test). Where instruction apparently had no effect, it can be argued that students were not focussed on form or were at least less focussed on form (Composition, 15 Minute Test, Sentence Test). The Term Examination also shows evidence of the effect of instruction, and it is likely that this test, like other standard tests, focussed testees on form by presenting words in isolation.

More Recent Data

More recent data on spelling is, in general, consistent with Rice and Cornman's results as well as with our hypotheses, but there are gaps.

Our hypotheses predict that when instruction has an effect, tests similar to the Column Test are used, tests that focus the testee on form. Several studies done since 1902 confirm that instruction can have an effect on spelling test performance. In some cases, the impact of instruction is substantial: Thompson (1930) found that instruction resulted in "permanent improvement equivalent to one-half year's gain in excess of that to be expected as a function of general maturity and incidental learning" (p. 71). In other cases, the impact is very modest: Guiles (1943) reported that children did only slightly better (4%) on a list of words whose spelling they had studied, compared to words they had not studied. In Fox and Eaton (1946), more frequent spelling instruction was related to better achievement (those who had instruction every day outperformed those who had spelling instruction four days per week), but differences were small; also, more daily time devoted to spelling instruction did not result in better spelling. In Thompson and in Fox

and Eaton's studies, spelling tests clearly focussed students on form (words were pronounced one at a time in isolation, then read in a sentence, then pronounced in isolation again). We have few details, however, on what kinds of tests were used in Guiles' study.

Both Rice and Cornman found that spelling accuracy in compositions was very high. This result has been replicated since then in numerous studies, summarized in Table 23. The statistics in this table are crude estimates: sample sizes, scoring procedures, and background characteristics of students varied. Nevertheless, the percentage of words spelled correctly is, in all cases, high, and accuracy tends to rise with age.

Considering that older writers use a wider variety of words (Jones, 1913, cited in Thompson, 1930) as well as more uncommon words (Haswell, 1988), data showing even slightly increasing spelling accuracy in compositions as children get older may actually reflect a profound increase in spelling ability.¹²

Table 23 Studies of spelling accuracy

study	description	grades	% of words spelled correctly
Johnson, 1917	impromptu essays	high school freshmen college freshmen	98.7% 99.4%
Ashbaugh, 1919	personal letters	7 9 12	97.4% 98.3% 98.6%
Brandenberg, 1919	essay exams (psychology)	college freshmen college sophomores college juniors college seniors	99.6% 99.5% 99.3% 99.6%
Lester, 1922	College Board Essays	high school seniors	99.8%
Hilderbrant, 1924	in-class writing	high school	98.9%
Lange, 1948	take-home essay (psychology)	college	98.9%
Kitzhaber, 1963	in-class essay "commentaries" journal entries	college freshmen college sophomores college seniors	99.8% 99.7% 99.6%
Pitts 7 Hirshfield, 1987 ^a	in-class essay "basic skills"	college age 9	97% 92.0%
Applebee,	narrative		

Langer, & Mullis, 1987	age 13 age 17	96.6% 97.6%
Wilde, 1988	in-class writing grade 3 grade 4	84.6% 87.6%
Clarke, 1988	in-class writing first grade invented spelling traditional ^b	58.4% 94%
Haswell, 1988	50" impromptu essay on assigned topic college freshmen college sophomores college juniors postgraduate ^c	97.9% 98.2% 98.3% 99.4%
Conners & Lunsford, 1988	essays written for writing class college freshmen & sophomores	99.6%

Notes:

Kitzhaber (1963) described by Haswell (1988).

a: 28 out of 71 subjects spoke English as a second language.

b: "traditional" students were focussed on correct spelling and were given extensive spelling help both before and during their writing. "Invented" spelling students received no such help but outperformed traditional spellers on two out of three spelling tests at the end of grade 1 (with no difference on the third test).

c: "postgraduate" subjects were "30-year-old or older employees in business, industry, and civil service deemed by their supervisors to be 'competent' writers in their workplace."

The reality of subconscious acquisition of spelling has been confirmed many times. Several studies have been published indicating that some children develop considerable competence in spelling before coming to school and without explicit instruction (e.g.

Goodman and Goodman, 1982). There is also evidence that children can spell many words they have not yet been taught (Curtiss and Dolch, 1939; Thompson, 1930), and that children who are excused from spelling instruction continue to improve (Kyte, 1948). In addition, it has been demonstrated that readers can improve their spelling of unfamiliar words from reading them in context (for experimental evidence, see Gilbert, 1934a, 1934b, 1935; Ormrod, 1986; Tyler, 1939, is a classroom study).¹³

The closest we have to a replication of Cornman's study of the effect of dropping spelling instruction is Richards (1920), who studied 78 children in grades 6, 7, and 8 who went without spelling instruction for one year. Richards reported that 67.5% of the children improved more than one year in spelling, while 20.4% made no change, and only 12% got worse. A standardized spelling examination was used.

Finally, Hammill, Larsen, and McNutt, (1977), in a cross-sectional study using the same test at different grade levels, compared children who had no spelling instruction with children who had formal instruction; the instructed children spelled better in grades 3 and 4, but the uninstructed children in grades 5 and 6 did just as well as the instructed children; the spelling test used apparently contained many of the words that children acquire through reading by grades 5 and 6.¹⁴

Conclusion

Our re-analysis confirms that Rice and Cornman's claims, for the most part, are correct: Spelling proficiency did not, in general, correlate with the amount of time devoted to spelling, and dropping spelling instruction typically had no significant effect on spelling. When instruction had an effect, the tests that were used focussed

students on form, and, we hypothesize, encouraged the use of conscious knowledge of spelling. Research done since Rice and Cornman confirms that incidental learning, or "acquisition," of spelling occurs.

If both acquisition and learning of spelling take place, and if learning is limited, we must ask, once again, the same question that prompted Rice's and Cornman's studies: Is direct teaching of spelling worth the trouble? Our suspicion is that it is not. Instruction may be only a modest help when students take word-list tests, tests that contain many words that will be acquired anyway with more reading. Perhaps the only spelling instruction well-read readers need is in the use of a spelling dictionary, to fill in the few gaps left by reading.¹⁵

FOOTNOTES

1. In some cases, Rice presents two scores for each school, an "a" score and a "b" score; for analyses presented in this paper, we used the "a" score, which represents children tested in the second semester of their grade level.
2. When small samples are involved, strict adherence to the .05 level may lead to type II errors. See e.g. Ferguson (1971, pp. 149-150) and Tate (1965, p. 227) for discussion.
3. Rice reported a number of additional findings, but supported his generalizations only with average (median?) scores. The averages are so close, however, that his conclusions would probably be supported by statistical analyses if the raw data were available. Specifically, Rice reported that children "of foreign parentage," children "hearing a foreign language at home," and children "of unskilled labors" scored practically the same as the overall average of the entire sample at all grade levels on both the Column Test and the Sentence Test. Rice also drew some conclusions for which no data was presented. He claimed that "some of the best results had been obtained where the phonic method had been employed" (p. 413), that better spelling was not associated with more general reading (p. 414), "nor did the schools where most time was devoted to written language make the best showing" (p. 414), nor was better spelling associated with the use of "oral spelling" (p. 413), "the flash method" (p. 414), or with an overall "mechanical" or "progressive" approach (p. 170).
4. We did not include data from Comman's school B in his Table 19 in our analysis, since only boys' scores were given. Also, for this part of the analysis, where boys' and girls' scores are given separately, we calculated and used the mean score. Finally, several scores in Comman's Table 19 were difficult to interpret (grade 5, schools E and K) and were thus omitted.
5. In computing simple correlations, we included the two experimental schools, Northwest and Agnew. As noted earlier, there was no spelling instruction given at these schools at the time this data was gathered (the Term Examinations were given in January, 1898; spelling instruction was dropped in Northwest and Agnew in September, 1897). We also computed simple correlations without Northwest and Agnew. None reached statistical significance, but all are positive:

grade	n	r
3	8	.552
4	8	.545
5	7	.376
6	9	.320
7	9	.160

- Excluding Northwest and Agnew results in a stronger case for instruction. Note that excluding these schools is a closer replication of Rice's study, since Rice did not include any schools in which spelling instruction was dropped. Data from Northwest and Agnew was not included in the calculation of cumulative correlations reported in our Table 7, since no record was provided of the amount of spelling instruction done in these schools before September, 1897.
6. As shown in Table 9, for the Northwest girls, $F = 1.983$, $p = .169$, not significant, but enough to make a conservative researcher pause. Examination of means shows practically no difference among the groups, however (6/96 comparison group = 96.22; 6/97 comparison group = 95.43; experimental group = 96.14) and post-hoc comparisons (Scheffe) were not significant.
 7. Comman also provided data on spelling accuracy for compositions written on topics in geography, science, history, and literature. We did not reanalyze this data, but present only essays Comman labelled "composition."
 8. Longitudinal analysis could not be applied to all the data given by Comman, since "a" and "b" semesters were combined in some cases.
 9. Comman's Table 25 contains additional data labelled "Column Test." Unfortunately, different words were used at each administration of the test, making comparisons impossible.
 10. Scores from Rice's subjects in Table 14 were computed by Comman from Rice's Table 2. These scores differ slightly from those in our Table 1, because we only included scores from schools that reported time devoted to spelling instruction, a subset of all the scores Rice presented.
 11. To get an idea of the impact of one year's instruction on test scores, we regressed Column Test and Term Examination scores on time devoted to spelling for each grade. For the Column Test, both simple and cumulative effects are presented, while for the Term Examination only cumulative effects are presented, because simple effects were very weak. (One could argue that simple correlations and regressions are more valid for the Term Examination, since it was a standard test given yearly - teachers may have taught to the test.)

Impact of Instruction:
Column Test: simple effects

grade	points gained with	
	30"/day	15"/day
4	20.58	10.29
5	17.34	8.67
6	10.11	5.06
7	6.06	3.03
8	5.55	2.78

Column Test: cumulative effects

grade	points gained with	
	30"/day	15"/day
5	14.40	7.20
6	18.72	9.36
7	6.60	3.30
8	9.60	4.80

Term Examination: cumulative effects

grade	points gained with	
	30"/day	15"/day
4	15.00	7.50
5	12.16	6.08
6	14.16	7.08
7	9.46	4.73

Even if only 15 minutes per day are devoted to spelling, as was the case in Cornman's schools (actually, 14.36 minutes per day), the regressions predict that instruction will have a clear effect, especially in the lower grades, assuming a linear relationship between time devoted to spelling and gains in spelling.

12. Of course, some people are poor spellers and make a lot of spelling mistakes. Our hypotheses predict that poor spellers need more reading, not more instruction, and that people who have read a lot are nearly always good spellers (but not necessarily perfect spellers; for discussion, see Krashen, 1985b, 1989).

13. Nisbet (1941) also found evidence for the existence of implicit learning of spelling through reading, but was not impressed with its effects. In his study, children ages 11 to 14 read and studied passages containing words they did not spell correctly on a pretest. After reading the passage, they could spell about one out of 25 target words. Nisbet concluded that "intensive reading and study of a passage then does lead to some learning of spelling, but this gain is not sufficient ... to justify the neglect of specific spelling instruction" (p. 11). If Nisbet's subjects read the passage only once, however, their progress was impressive. Given

enough reading, this small gain may be enough to account for a great deal, if not all, of spelling competence. Nagy, Herman, and Anderson (1985) have made this argument for vocabulary development. They estimated that a reader's chance of acquiring a new word from a single exposure was between five and 20%, depending on the test used. According to their calculations, if children read about one million words per year, this small gain accounts for the vocabulary growth observed in children, about several thousand words per year. (See also Ormrod, 1986, for results and interpretation similar to Nisbet's.)

Nisbet also briefly reported on a study in which he found that looking up words in a dictionary greatly facilitated learning how to spell them: "... the average effect of looking up a word, even if only to find the meaning, is so great that if a pupil could not spell the word before, the chances are one in two that he will be able to spell it afterwards" (p. 11). Nisbet presents no data to support this (his report is a summary of his thesis), but even if this result is true it is not a strong argument for dictionary work. First, looking up words in the dictionary takes time, time that could be spent reading. Second, looking up words is boring. Many people would rather not read at all if they had to use the dictionary extensively. Free voluntary reading, on the other hand, is a highly pleasurable activity (Nell, 1988) that provides other benefits in addition to improved spelling.

14. It is interesting to note that instruction had its most consistent effect in grade 3 in Cornman's data. Instructed children, the comparison groups, were superior in seven out of eight comparisons on the Column Test, six out of eight on the Sentence Test, and on the Term Examination. Also, according to regression analysis (note 11), the impact of instruction on the Column Test and Sentence Test is greatest in lower grades.

15. It could be argued that our analysis shows only that traditional spelling instruction relying on word lists is of limited value, and that more enlightened methods of teaching spelling could be of greater value. While some techniques of teaching spelling are more efficient than others (see Fitzsimmons and Loomer, 1982, for a review), it appears to be the case that a great deal of spelling competence can be acquired.

REFERENCES

- APPLEBEE, A., LANGER, J., & MULLIS, I. (1987): *Grammar, punctuation, and spelling: Controlling the conventions of written English at ages 9, 13, and 17*. Princeton: Educational Testing Service.
- BRANDENBURG, G.C. (1919): Some possibly secondary factors in spelling ability. *School and Society*, 9, 632-636.
- CLARKE, L. (1988): Invented versus traditional spelling in first graders' writings: Effects on learning to spell and read. *Research in the Teaching of English*, 22, 281-310.
- CONNERS, R. & LUNSFORD, A. (1988): Frequency of formal errors in current college writing, or Ma and Pa Kettle do research. *College Composition and Communication*, 4, 395-409.
- CORNMAN, O. (1902): *Spelling in the elementary school*. Boston: Ginn.
- CURTISS, H. & DOLCH, E. (1939): Do spelling-books teach spelling? *Elementary School Journal*, 39, 584-592.
- FERGUSON, G. (1971): *Statistical analysis in psychology and education*. New York: McGraw-Hill.
- FTTZSIMMONS, R. & LOOMER, B. (1982): Excerpts from Spelling: Learning and instruction - research and practice. In W. Barbe, A. Francis, and L. Braun (Eds.), *Spelling: Basic skills for effective communication* (pp. 61-79). Columbus, Ohio: Zaner-Bloser, Inc.
- FOX, W. & EATON, M. (1946): Analysis of spelling proficiency in 82,833 pupils in grades 2 through 8 in 3,547 teaching units of the city schools in Indiana. *Bulletin of the School of Education, Indiana University*, 22, 9-45.
- GILBERT, L.C. (1934a): Effect of spelling on reading in the ninth grade. *School Review*, 42, 197-204.
- GILBERT, L.C. (1934b): Effects of reading on spelling in the secondary schools. *California Quarterly of Secondary Education*, 9, 269-75.
- GILBERT, L.C. (1935): Study of the effect of reading on spelling. *Journal of Educational Research*, 28, 570-576.
- GOODMAN, K. & GOODMAN, Y. (1982): Spelling ability of a self-taught reader. In F. Gollasch (Ed.), *Language and literacy: The selected writings of Kenneth S. Goodman*, vol. 2. (pp. 221-226). London: Routledge and Kegan Paul.
- GUILES, R. (1943): Effect of formal spelling on spelling accuracy. *Journal of Educational Research*, 37, 284-289.
- HAMMILL, D., LARSEN, S., & MCNUITT, G. (1977): The effect of spelling instruction: A preliminary study. *Elementary School Journal*, 78, 67-72.
- HASWELL, R. (1988): Error and change in college student writing. *Written Communication*, 5, 479-499.
- HILDERBRANT, E. (1924): Can high school students spell? *The School Review*, 32, 779-781.
- JOHNSON, R. (1917): The persistency of error in English composition. *School Review*, 25, 555-580.
- KRASHEN, S. (1981): *Second language acquisition and second language learning*. New York: Prentice-Hall.
- KRASHEN, S. (1982): *Principles and practice in second language acquisition*. New York: Prentice-Hall.
- KRASHEN, S. (1985a): *The input hypothesis: Issues and implications*. White Plains, New York: Longman.
- KRASHEN, S. (1985b): *Inquiries and insights*. Menlo Park, Ca.: Alemany Press.
- KRASHEN, S. (1990): The input hypothesis and some competing hypotheses. In R. Courchene and J. St. John (Eds.), *Comprehension-based language teaching: Current trends*. Ottawa: University of Ottawa Press.
- KYTE, G. (1948): When spelling has been mastered in the elementary school. *Journal of Educational Research*, 42, 47-53.
- LANGE, P. (1948): A sampling of composition errors of college freshmen in a course other than English. *Journal of Educational Research*, 42, 191-200.
- LESTER, J. (1922): A study of high school spelling material. *Journal of Educational Psychology*, 2, 65-74; 152-159.
- NAGY, W., HERMAN, P., & ANDERSON, R. (1985): Learning words from context. *Reading Research Quarterly*, 20, 233-253.
- NELL, V. (1988). The psychology of reading for pleasure: Needs and gratifications. *Reading Research Quarterly*, 23, 6-50.
- NISBET, S. (1941): The scientific investigation of spelling instruction: Two preliminary problems. *British Journal of Educational Psychology*, 11, 150.
- ORMROD, J. (1986): Learning to spell: Three studies at the university level. *Research in the Teaching of English*, 20, 160-173.
- PETERS, M. (1985): *Spelling: Caught or taught?* London: Routledge and Kegan Paul.
- PITTS, S. & HIRSCHFELD, G. (1987): Three spelling measures as correlates of reading ability in underprepared college freshmen. *Reading Horizons*, 27, 103-110.

- RICE, J.M. (1897): The utility of the spelling grind. *Forum*, 23, 163-172, 409-419.
- RICHARDS, A. (1920): Spelling and the individual system. *School and Society*, 10, 647-650.
- SMITH, F. (1981): Reading: What is basic? In M. Douglas (Ed.), *Claremont Reading Conference, 45th Yearbook* (pp. 1-20). Claremont, California: Claremont Graduate School of Education.
- TATE, M. (1965): *Statistics in education and psychology*. New York: Macmillan.
- THOMPSON, R. (1930): The effectiveness of modern spelling instruction. *Contributions to Education*, No. 436. New York: Teachers College, Columbia University.
- TIDYMAN, W.F. (1915): A critical study of Rice's investigation of spelling efficiency. *Pedagogical Seminary*, 22, 391-400.
- TYLER, I.K. (1939): Spelling as a secondary learning. *Contributions to Education*, No. 781. Teachers College, Columbia University.
- WALLIN, J.E.W. (1910): Has the drill become obsolescent? *Journal of Educational Psychology*, 1, 200-213.
- WILDE, S. (1988): Learning to spell and punctuate: A study of eight- and nine-year-old children. *Language and Education*, 2, 35-59.

la-Neuve, Belgium; Yorick Wilks, Department of Language and Linguistics, University of Essex, Colchester, England; Antonio Zampolli, Istituto del Consiglio Nazionale delle Ricerche, Pisa.

Articles are accepted for publication in the field of applied linguistics in a broad sense. Authors receive twenty offprints free of charge. The only accepted abbreviation for bibliographical purposes is I.T.L. The subscription price for the volume year, four issues, is 900 Belgian Francs. When purchased directly, paid by Postal Cheque Account 000-1020401-58 or Banking Account 431-0371811-77 individuals pay 900 BF and booksellers 960 BF. When paid through a foreign bank the subscription price is 1.300 BF. Single issues may be ordered at a price of 1.000 BF per issue. Prices are postage included. All correspondence should be addressed to the managing editor.

INSTRUCTIONS TO AUTHORS

Authors are requested to prepare their contributions in accordance with the following conventions :

Manuscript : The preferred languages of publication are English, French, German and Spanish. The manuscript, including footnotes and references, should be type-written, double-spaced, on consecutively numbered sheets.

References : The reference list should be alphabetically ordered and added on (a) separate sheet(s).

Illustrations : All illustrations (figures, tables, drawings) should be clearly drawn on separate sheets. At the appropriate place in the text an indication must be given of where illustrations have to be inserted. Authors are requested to send original copies only. We also accept floppy-disks (two-sided, double density, 5.25" and 3.5").

ISSN 0019-0810