The effect of adding supplementary writing to an extensive reading program

by Beniko Mason

Abstract

This study investigated whether adding supplementary writing to an extensive reading program would increase its effectiveness for the development of grammatical accuracy. The participants were Japanese female college learners of English ($N = 104$) studying in an extensive reading program. The Japanese summary group ($n = 34$) wrote summaries in Japanese, the English summary group ($n = 34$) wrote summaries in English, and the Correction group ($n = 36$) wrote summaries in English, received corrective feedback, and rewrote their corrected summaries. All participants read an average of 2300 pages (about 500,000 words) in three semesters, and the Correction group's summaries were corrected 25 times. The results revealed that all three groups improved significantly, and there were no statistically significant differences among the groups on three tests. The questionnaire revealed that the Japanese summary group spent 150 hours reading while the other groups spent about 300 hours reading, writing and rewriting. The conclusion was that adding supplementary writing did not lead to greater accuracy and that it was inefficient.

1. INTRODUCTION

The last few decades have witnessed the publication of a considerable amount of empirical evidence supporting what Krashen has called "the power of reading" (Krashen, 1993). There is little doubt that reading itself leads to better reading, better vocabulary, better writing, and better control of grammar in both first and second languages. The impact of reading has been demonstrated in controlled studies of in-school reading ("sustained silent reading," and "extensive reading"; see e.g. Elley and Mangubhai, 1983; Mason and Krashen, 1997), as well as in numerous case histories (e.g. Krashen, 1993; Cho and Krashen, 1994) and correlational studies of self-reported recreational reading (e.g. Anderson, Wilson, and Fielding, 1988).

While there is agreement that recreational reading is helpful, it can be asked whether reading should be supplemented with other activities in order to produce the best results. Can recreational reading be enhanced by the use of supplementary activities? Can we, in other words, increase the power of reading? A wide range of supplementary activities are possible, but the supplement that appears to be the most popular is to include writing that is related to what has been read.

It may come as a surprise to many readers, but there is no evidence that writing alone increases language or literacy proficiency, that is, increasing the amount of writing done does not increase proficiency. Reviews of first language studies can be found in Krashen (1993). Burger (1989) reported that adding an extra class on writing, which included correction of students' written errors, had no impact on gains in English proficiency on a variety of measures for adult students of ESL taking sheltered classes in Canada. Tsang (1996) reported that Hong Kong middle and high school students who participated in an after-school extensive reading program lasting 24 weeks made better gains in writing than comparison students who did extra writing rather than reading. Not yet investigated, however, is whether a program integrating reading and writing will be more efficient and effective than reading alone.

It can be argued that writing alone is insufficient, that writing requires feedback on form, that is, grammar correction, to be effective. Once again, the research is discouraging. Several reviews have concluded that the impact of correction is very limited: In many cases, there is no impact at all on accuracy, and when an effect is present it is very modest and confined to situations in which students are heavily focused on form (Truscott, 1996; Krashen, 2002). Not yet investigated, however, is whether grammar correction on student written output can enhance the impact of reading.

It is important to continue to investigate the impact of output and grammar correction, despite the lack of supporting empirical evidence so far, and to continue to see under what conditions they might be effective. It is nearly an unquestioned assumption that "we learn to write by writing" and many students request correction of form (e.g. Cathcart and Olsen, 1976).

The goal of this study is thus to compare the impact of extensive reading with three kinds of supplementation: students writing summaries in their primary language (Japanese) of what they have read, a condition that relies only on reading for language development; students writing summaries in English, a test of the hypothesis that supplementation using writing will enhance the power of reading, and students writing summaries in English, having their errors corrected, and rewriting the summaries. The rewriting condition was included because of claims that correction alone is insufficient: It has been claimed that students must also rewrite and incorporate the corrections in a subsequent version of their paper (Chandler, 2003). This third condition tests the hypothesis that additional writing plus grammar correction will enhance the power of reading.

Because of the claim that error correction might have different effects on different measures, three different tests were used, including one that allowed a considerable amount of focus on form under conditions similar to those present during the treatment, grammatical accuracy in writing.

An interesting feature of the design was that it was possible to ensure that students were in agreement with each method of supplementation; those who wrote English summaries agreed that this was an effective plan, and those who had their errors corrected were unanimous in their desire to receive grammar correction. In a sense, this loaded the study for success, reducing the chance that a negative result was due to students' discomfort with the method used.

2. METHOD

2.1 Subjects

The participants in this study were 104 first-year female English majors in an extensive reading (ER) class at a junior college in Osaka, Japan. Their average scores on the Test of English for International Communication (TOEIC) were 123.64 ($SD = 35.19$) for the reading section and 153.62 ($SD = 47.51$) for the listening section out of a possible score of 495 for each section. The TOEFL equivalent of the total score would be approximately 351-371 (Axe & Belle, 2004).

2.2 Treatments

All participants were enrolled in eight classes per week. Six out of the eight were identical; all focused on listening and speaking, all were taught in English by native speakers of English, and all used the same textbooks and the same audio tapes. Subjects also were enrolled in either a grammar or phonology class that was taught in Japanese. The eighth class was the Extensive Reading (ER) class which was held once a week.

ER students were asked to read 1,000 pages (about 250,000 words) from graded readers each semester. At the beginning of the study, subjects read an orientation booklet written by this researcher and viewed a video that explained the program. Almost 100% of the students stated that they understood the significance of the program and would attempt to do the reading. Approximately 5000 books were arranged according to level in the ER classroom, and the students selected books weekly according to their proficiency level and interest. All participants started reading beginning level graded readers (600 word level) and gradually moved up to higher level readers (1100, 1600, and 2200 word levels) and to authentic books written for young native-speaking adults.

Reading was done mostly at home. Students were required to keep a record of their reading homework, including the number of pages read, and to submit a notebook every week in which they wrote a brief summary for each book they read. The normal procedure was to require that this summary be written in English. They also wrote their reflections on the content of what they read and wrote comments about their progress in reading in Japanese.

After two weeks, listening to stories was incorporated into the extensive reading class. Listening to stories provides comprehensible input for the development of overall language competence, including listening and vocabulary (Allen & Allen, 1985; Elley, 1989; Brett, Rothlein, & Hurley, 1996; Vivas, 1996). The classroom procedure and the stories were the same for all the classes.

The Extensive Reading classes formed themselves into three experimental groups in the following manner: A few weeks after the beginning of the ER class in April, the students in one class (henceforth class JSG, Japanese Summary Group) requested that they write their summaries in Japanese rather than English, because they felt that it was too difficult to write in English. Another class (henceforth class CORRECTION) requested that their English summaries be corrected. Other classes did not request anything beyond the required work. A third class (henceforth class ESG, English Summary Group) was selected to be a third experimental group that would write their summaries in English. Class ESG was chosen because classes JSG, COR and ESG all met in the afternoon. Thus, the three groups were: (a) a group that read extensively and wrote book summaries in Japanese (JSG: Japanese Summary Group, $n = 34$), (b) a group that read extensively and wrote book summaries in English (ESG: English Summary Group, $n = 34$), and (c) a group that read extensively, wrote book summaries in English, received corrective feedback from a native speaker of English, rewrote the corrected summaries, and submitted the rewritten summary (Correction Group, $n = 36$). Thus, participants in this study chose their treatment, and were not forced to do anything that they did not agree to do.

An experienced teacher, a native speaker of British English with a master's degree in Second Language Acquisition provided corrective feedback on the summaries written by the correction group. He had been at this junior college for over ten years. We agreed on the following points regarding the feedback. He would:

(a)   concentrate on global errors that affect overall meaning and organization,

(b)   mark the error and sometimes supply the correct form, and sometimes not, using his own judgment as to whether it was necessary to provide the form,

(c)   indicate when he did not understand the story line,

(d)   note whether the story was coherent or complete,

(e)   point out grammatical errors that he feels are necessary for the learner to pay attention to, but

(f)   not correct every grammatical error.

Errors were corrected 25 times over three semesters for participants in the Correction group who submitted summaries.

2.3 Reliability of Corrective Feedback

To investigate the consistency and the systematicity of the corrections, summaries written by first year female students at the same junior college ($N = 29$) who were not in this study were used. The students' English proficiency was approximately the same as the participants in this study. The instructor was asked to correct their summaries as he had corrected those written by the participants in the study.

The students were first provided with several different graded readers at the 600 word level, which was considered to be easily within their reading competence, and were asked to choose one to read. All consisted of a short story of about 2000 to 2500 words in length. Students read for about 15 to 20 minutes. They then wrote a summary of the story they read in English. Students had plenty of time to write the summary, about 90 minutes including the reading time. These conditions were thus similar to the conditions used in the actual study.

The instructor corrected the same papers twice, with the second grading occurring one month after the first. As shown in table 1, the instructor corrected fewer errors the first time (267) than in the second (470). The kinds of corrections made, however, were very similar, confirming that correction was consistent (for additional details, see Mason, 2003).

Table 1. *Types of Correction at Two Different Times*

	Dec. 1999	Jan. 2000
Types of correction	Frequency (%)	Frequency (%)
Spelling	77 (29%)	88 (19%)
Articles	48 (18%)	77 (16%)
Tense	48 (18%)	58 (12%)
Prepositions	21 (8%)	40 (9%)
Misuse of Words	16 (6%)	33 (7%)
Infinitive	16 (6%)	23 (5%)
Plural	8 (3%)	23 (5%)
Other grammar forms	33 (12%)	128 (27%)
Total	267 (100%)	470 (100%)

2.4 Measures

The measures used were a 100-item cloze test (test-retest reliability = .87), the reading section of the TOEIC (Test of English for International Communication) test (KR21 = .96), and the number of error free clauses made per 100 words in writing (inter-rater r = .90). The same measures were used at the beginning and end of the study; the cloze passage used was identical, but an alternate form of the TOEIC was used, and the prompt (story) for the pre and post writing samples differed.

To generate a writing sample for the error-free clause calculation, students were asked to read a short story, and write a summary in English. The pretest prompt was taken from a story at the 600-word level and the posttest prompt was from an un-graded text. The following is an excerpt from the un-graded text,

Poor little Lisa, how she cried! All she had left were the little red hood, and her pretty red shoes. She had given away all her other things. She had even lost her basket! She must have forgotten to pick it up when the bear frightened her. As the shadows grew darker and darker in the big black forest, she tried hard to be brave and to keep on walking. The stars twinkled brightly in the black sky while the Old Man in the Moon smiled kindly down on her. Tired out, she sat down on a big stone to rest.

("Little Lisa," Nerman, 1955)

Students were asked how much time it took them to read the prompt after the post-test; the average was about 30 minutes, and analysis of variance revealed no significant difference among the three groups. Students were also asked how many pages they had read at the end of each semester, and how much time they had spent reading and writing English summaries. At the end of the third semester, students were asked a) Do you think your writing ability improved? b) Did summary writing assignment hinder your reading? c) Did you sometimes copy from a book when you wrote a summary? d) How much did you copy (0%, 5%, 25%, 50% and 90%)? e) Should we continue writing summaries in English? f) Was writing summaries in English more tiring than reading?

The research hypotheses for this study were that there would be no statistically significant differences among the groups on the mean score of the cloze posttest, the reading section of the TOEIC posttest, and on error free clauses per 100 words on the writing posttest. The alpha level was set at .01, as multiple ANOVAs were used for the analyses. Threats to validity such as maturation exist in longitudinal research, but this research was continued for three semesters because error correction feedback might need a long amount of time to have an effect on grammatical accuracy in writing (Franzen, 1995; Rod Ellis, personal communication, 1999).

3. RESULTS

3.1 Cloze test

Table 2 presents descriptive statistics. A one-way ANOVA showed that there was no significant difference among the groups on the pretest, $F(2, 93) = 1.514$, $p = .225$. A tests (pretest/posttest) by groups (three levels) repeated-measures two-way analysis of variance showed that there was a statistically significant difference within the participants, $F(1, 91) = 359.274$, $p = .000$, indicating that all groups improved significantly across the pretest and the posttest. No statistically significant differences were found among the groups, $F(2, 91) = 1.909$, $p = .154$. In addition, the interaction was not significant, $F(2, 91) = .549$, $p = .580$ (Table 3).

Table 2. *Descriptive Statistics for Cloze Tests*

	Pretest	*n*	Posttest	*n*	Gain
Group	*M (SD)*		*M (SD)*		
JSG	30.63(7.70)	31	45.43(6.90)	32	14.80
ESG	28.42(8.07)	29	42.42(6.60)	33	14.00
Correction	27.00(8.56)	36	42.97(8.04)	36	15.97

JSG = Japanese summary group
ESG = English summary group

Table 3. *Repeated-Measures Two-Way Analysis of Variance on the Cloze Test*

	df	*SS*	*MS*	*F*	*p*
Between	2	338.82	169.41	1.91	.15
Within	1	10351.37	10351.37	359.27	.00
Interaction	2	31.61	15.80	.55	.58
Total	5	10721.80			

3.2 TOEIC

Table 4 presents descriptive statistics. A one-way ANOVA showed that there was no statistically significant difference among the groups on the TOEIC pretest, $F(2, 94) = 1.82$, $p = .17$. A tests (two tests) by groups (three levels) repeated-measures two-way ANOVA revealed a statistically significant difference within the participants, $F(1, 85) = 53.71$, $p = 00$, showing that all groups gained, but there was no statistically significant difference among the groups, $F(2, 85) = 2.25$, $p = .11$, and the interaction was not significant. $F(2, 85) = .20$, $p = .82$ (Table 5).

Table 4. *Descriptive Statistics for the Reading Section of the TOEIC Pretest and Posttest*

	Pretest	*n*	Posttest	*n*	Gain
Group	*M (SD)*		*M (SD)*		
JSG	129.83(33.04)	30	163.50(38.03)	30	33.67
ESG	112.66(36.42)	30	146.83(43.77)	30	34.17
Correction	121.78(25.06)	28	162.32(53.99)	28	40.54

JSG = Japanese summary group
ESG = English summary group

Table 5. *Repeated-Measures Two-Way Analysis for Variance on the TOEIC*

	df	*SS*	*MS*	*F*	*p*
Between	2	9151.76	4574.88	2.25	.11
Within	1	57353.69	57353.69	53.71	.00
Interaction	2	420.08	210.04	.20	.82
Total	5	66925.58			

3.3 Error Free Clause Test

In response to the question asked as the writing test, the average time spent for reading the prompt was about 30 minutes, and as mentioned earlier, analysis of variance revealed no significant difference among the three groups. This shows that the participants spent about 30 minutes for reading and 30 minutes for writing.

Table 6 presents descriptive statistics for the number of error free clauses written in 100 words.

Table 6. *Descriptive Statistics for the Error Free Clause Ratio Data*

	Pretest	*n*	Posttest	*n*	Gain
Group	*M (SD)*		*M (SD)*		
JSG	8.95 (3.56)	34	12.19(2.67)	32	3.24
ESG	8.05 (4.25)	34	10.37(2.87)	33	2.32
Correction	9.62 (3.26)	36	11.30(1.87)	36	1.67

JSG = Japanese summary group
ESG = English summary group

A one-way ANOVA showed no statistically significant differences among the groups, $F(2,101) = 1.58$, $p = .21$ on the number of error free clauses per 100 words on the pretest. A tests (pretest/posttest) by groups (three levels) repeated-measures ANOVA showed that there was a statistically significant difference within the participants ($p = .00$), but no statistically significant difference among the groups ($p = .05$) (Table 7). All groups improved to the same degree.

Table 7. *A Repeated-Measures Two-Way Analysis of Variance on the EFC per 100 words Ratio Data*

	df	*SS*	*MS*	*F*	*p*
Between	2	78.702	39.352	3.193	.05
Within	1	291.829	291.829	38.334	.00
Interaction	2	19.618	9.808	1.288	.28
Total	5	390.149			

3.4 Effect Sizes

Effect sizes were calculated using a method that takes pretest scores into account by subtracting the effect size of the comparison group gain from the effect size of the experimental group gain, using pooled standard deviations (Rudner, Glass, Evartt, & Emery, 2002). Table 8 presents the effect sizes between the groups. Note that the group that wrote in Japanese was superior to all other groups in five out of six comparisons. (A plus sign indicates that the first group of each pair was superior.)

Table 8. *Effect Sizes*

		Comparison Pairs	
Test	JSG: ESG	ESG: Correction	JSG: Correction
Cloze	+0.12	- 0.02	+ 0.15
Reading/TOEIC	+ 0.10	- 0.17	- 0.08
EFC	+0.39	0.00	+ 0.39

JSG= Japanese summary group
ESG = English summary group

3.5 Pages Read

The groups read about the same number of pages during the first and second semesters, but in the third semester, the Correction group read more than the other groups (Table 9).

Table 9. *Descriptive Statistics for Number of Pages Read*

			Semester		Total
Group		1st	2nd	3rd	
JSG					
	M	965.09	591.67	663.53	2220.29
	SD	364.32	295.95	283.44	
	n	33	34	32	
ESG					
	M	912.86	459.76	871.84	2244.46
	SD	154.95	266.98	214.98	
	n	33	30	31	
Correction					
	M	954.06	572.42	941.14	2467.61
	SD	156.78	199.20	140.23	
	n	36	36	36	

JSG = Japanese summary group
ESG = English summary group

3.6 Time Spent for Reading

The participants reported approximately how much time they spent reading per week. The Japanese summary writing group spent the least amount of time reading books in English (Table 10), and the Correction group devoted the most time to reading. This result is consistent with results regarding the number of pages read.

Table 10. *Descriptive Statistics for Hours Spent on Reading per Week and Total Hours*

Group	M	SD	n	Total
JSG	3.35	1.01	30	150.75
ESG	3.71	1.38	21	167.13
Correction	4.44	1.57	35	199.94

JSG = Japanese summary group
ESG = English summary group

3.7 Time Spent for Writing

The English summary group reported that they spent about 2.2 hours per week the first and second semesters and 1.78 hours per week the third semester writing summaries. Participants in the Correction group reported that they spent about 2 hours per week the first and second semester and 2.53 hours per week the third semester doing the same task (Table 11). If their reports are accurate, multiplying the number of hours by the total number of weeks in the three semesters (45 weeks) means that both the English Summary and Correction groups spent about 100 hours writing and rewriting in English, while the Japanese summary writing group spent no time writing in English.

Table 11. *Reported Hours Spent per Week Writing Summaries in English*

		Semester	
Group	1st & 2nd (n)	3rd (n)	Total
ESG	2.20 (20)	1.78	92.75 (45 weeks)
Correction	1.99 (35)	2.53	97.58 (45 weeks)

ESG = English summary group

The Correction group spent a total of about 300 hours (297.51) reading and writing and the English summary group spent about 260 hours (259.98), while the Japanese summary group spent about 151 hours (150.750) reading and no time writing in English.

3.8 Efficiency per Hour

 Table 12 presents the relative efficiency of the three groups. In each case, the gain scores were divided by the number of total hours spent for English study. The JSG was about twice as efficient as the other two groups on all three measures.

Table 12. *Efficiency per Hour*

Group	Cloze	TOEIC Reading	Error Free Clause
JSG (151 hours)	.098	.223	.021
ESG (260 hours)	.054	.131	.008
Correction (298 hours)	.054	.136	.006

3.9 Response to Questions

In response to the questions asked at the end of the study (Table 13), two-thirds of the participants in the Japanese summary and the Correction groups felt that they had improved in writing, and more than half of the participants in the English summary group also felt that they improved (Question 1). However, nearly half of the participants in both English summary group and Correction group felt that summary assignment in English hindered their reading (Question 2) and about half felt that it was more tiring than reading (Question 6). Furthermore, two-thirds of the participants in English and Correction groups said that they had copied from a book when they wrote a summary (Question 3), and a significant amount of the work written by the participants in both English summary and Correction groups was not their own (Question 4). Nevertheless, 65% of the participants from English summary and Correction groups believed that they should write summaries in English after reading (Question 5).

Table 13. *Responses to the questions*

(1) Do you think your writing ability improved?

Group	Yes	No	I don't know	No Response	Total
JSG	25(74%)	9(26%)	-		34(100%)
ESG	22(58%)	11(29%)	-	5(13%)	38(100%
Correction	25(74%)	7(21%)	-	2(5%)	34(100%)

(2) Did summary writing assignment hinder your reading?

Group	Yes	No	I don't know	No Response	Total
JSG	14(42%)	20(58%)	-		34(100%
ESG	16(42%)	17(45%)	-	5(13%)	38(100%)
Correction	18(53%)	13(38%)	1(3%)	2(6%)	34(100%)

(3) Did you sometimes copy from a book when you wrote a summary?

Group	Yes	No	I don't know	No Response	Total
ESG	25(66%)	7(18%)	-	6(16%)	38(100%)
Correction	22(65%)	9(26%)	-	3(9%)	34(100%)

(4) How much did you copy?

Group	0%	5%	25%	50%	90%	No Response	Total
ESG	0	20(53%)	13(34%)	1(3%)	0	4(10%)	38(100%)
Correction	5(15%)	14(41%)	11'32%)	1(3%)	0	3(9%)	34(199%)

(5) Should we continue writing summaries in English?

Group	Yes, necessary	No, Not Necessary	Total
ESG	24(63%)	14(37%)	38(100%)
Correction	22(65%)	12(35%)	34(100%)

(6) Was writing summaries in English more tiring than reading?

Group	Yes	No	No Response	Total
ESG	19(50%)	13(34%)	6(16%)	38(100%
Correction	19(56%)	13(38%)	2(6%)	34(100%)

JSG = Japanese summary group
ESG = English summary group

It was observed that the rate of summary submission for the Correction group decreased from 60% to 30% towards the end of the study even though all students in the Correction group agreed to correct their papers and hand in revisions. The rate of summary submission for the other two groups also decreased about the same amount.

4. DISCUSSION

All three groups in this study improved significantly, but there were no significant differences among the groups in gains. The group that wrote summaries in Japanese, their first language, was the most efficient, making the greatest gains in terms of points gained for the time devoted to English.

As noted earlier, those who wrote English summaries said that they desired this kind of supplementation before the treatment began. When the treatment was over, most students in these groups still felt that summary writing was helpful (table 12). But a large percentage felt that writing summaries hindered their reading, that it was tiring, and about two-thirds of the English summary writers admitted that they sometimes copied part of their summary. All groups, including those who wrote summaries in Japanese, handed in fewer summaries as the treatment progressed.

These results do not definitely demonstrate that output and output plus correction are always ineffective. It is of course possible that there simply wasn't enough output or correction or that the means employed were not optimal.

It may be the case that output must be "comprehensible output" (Swain, 1985; but see Krashen, 2003), done in a way to encourage feedback on meaning. According to the Comprehensible Output hypothesis, output helps language acquisition when listeners or readers fail to understand the message, forcing the language acquirer to try again, with an improved version. This kind of communicative pressure was not employed in this study. It may be the case that the correction done needed to be more "selective, prioritized, and clear" (Ferris, 1999; but see Truscott, 1999), that is, focused on certain rules, with some corrections given higher priority, and done in a way that makes it obvious what the problem is and what needs to be done to repair the error. In this study, correction was, at times, confined to only pointing out that an error was made, and at other times included the correct form. Correction was quite consistent, but was not directed at certain points of grammar to the exclusion of other points.

What we can conclude, however, is that output in the form of summary writing, with and without the kind of correction usually provided in language classes, did not add to the power of reading.

Insisting that output must be accompanied by feedback on communicative success, and/or that error correction needs to be done in a precise manner is equivalent to saying that extraordinary efforts on the parts of students and teachers are necessary to improve on the power of reading, a conclusion that leaves unexplained the fact that so many students have acquired significant amounts of language without them.

REFERENCES

Allen, V. G., Allen, E. D., 1985. Story retelling: Developmental stages in second-language acquisition. The Canadian Modern Language Review, 41, 686-691.

Anderson, R.C., Wilson, P.T., Fielding, L.G., 1988. Growth in reading and how children spend their time outside of school. Reading Research Quarterly, 23, 285- 303.

Axe, R., Bell, I., 2004. Using TOEFL ITP and TOEIC to assess students. Paper presented at the 38th annual TESOL convention, Long Beach, CA.

Brett, A., Rothlein, L., Hurley, M., 1996. Vocabulary acquisition from listening to stories and explanations of target words. The Elementary School Journal. 96, 415-421.

Burger, S., 1989. Content-Based ESL in a Sheltered Psychology Course: Input, Output, and Outcomes. TESL Canada, 6, 45-59.

Cathcart, R., Olsen, J., 1976. Teachers' and students' preferences for correction of classroom conversation errors. In J. Fanselow and R. Crymes (Eds.) On TESOL '76. Washington, DC: TESOL, 41-53.

Chandler, J., 2003. The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. Journal of Second Language Writing, 12(3), 267-296.

Cho. K. S., Krashen, S., 1994. Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. Journal of Reading, 37, 662-667.

Elley, W. B., 1989. Vocabulary acquisition from listening to stories. Reading Research Quarterly, 24, 174-187.

Elley, W. B., Mangubhai, F., 1983. The impact of reading on second language learning. Reading Research Quarterly, 19, 53-67.

Ferris, D., 1999. The case for grammar correction in L2 writing classes: a response to Truscott (1996). Journal of Second Language Writing 8(1): 1-11.

Frantzen, D., 1995. The effects of grammar supplementation on writing accuracy in an intermediate Spanish content course. The Modern Language Journal, 79, 331- 344.

Gradman, H. L., Hanania, E., 1991. Language learning background factors and ESL proficiency. Modern Language Journal, 75, 39-50.

Krashen, S., 1993. The power of reading. Libraries Unlimited, Englewood, CO.

Krashen, S., 2002. The comprehension hypothesis and its rivals. In: Selected papers from the Eleventh International Symposium on English Teaching/Fourth Pan- Asian Conference. English Teachers Association/ROC, 395-404.

Krashen, S. D., 2203. Explorations in language acquisition and use. Heinemann, Portsmouth, NH.

Mason, B., 2003. A study of extensive reading and the development of grammatical accuracy by Japanese university students learning English. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.

Mason, B., Krashen, S., 1997. Extensive reading in English as a foreign language. System, 25, 91-102.

Nerman, E., 1955. Little Lisa. In: Martignoni, M. E. (Ed.). The Illustrated Treasury of Children's Literature. Grosset & Dunlap Inc., New York, NY, pp. 282-284.

Runder, L., Glass, G., Evartt, D., Emery, P., 2002. A user's guide to the meta-analysis of research studies. ERIC Clearinghouse on Assessment and Evaluation, College Park, MD.

Swain, M., 1985. Communicative competence: some roles of comprehensible input and comprehensible output in its development. In Gass, G., Madden, C. (Eds.), Input in second language acquisition. Newbury House, New York, NY, pp. 235- 254.

Truscott, J., 1996. The case against grammar correction in L2 writing classes. Language Learning, 46, 327-369.

Truscott, J., 1999. What's wrong with oral grammar correction? The Canadian Modern Language Review, 55, 437-455.

Tsang, W. K., 1996. Comparing the effects of reading and writing on writing performance. Applied Linguistics, 17, 210-233.

Vivas, E., 1996. Effects of story reading on language. Language Learning, 46, 189-216.