Seeking a Role for Grammar: A Review of Some Recent Studies

Stephen D. Krashen
University of Southern California

ABSTRACT Studies of the impact of formal instruction consistently show that more instruction results in, at best, modest increases in consciously-learned competence, a conclusion that is consistent with the claims of the Monitor hypothesis.

Introduction

In Krashen (1992, 1993) it was argued that studies that attempted to demonstrate the efficacy of direct instruction in grammar showed only that grammar teaching has a peripheral effect. In this paper, I review studies that have appeared since that time, and come to the same conclusion, in disagreement with the researchers themselves, who in nearly every case conclude that they have shown that grammar instruction and focusing on form "works."

I define direct grammar instruction here as consisting of two components: (1) focus on form, and (2) presentation of the rule. It is possible to do (1) without (2), that is, focus students on form without presenting a rule. This is done in several studies described below. It is not, however, possible to do (2) without (1): When we present a rule, we are also focusing students on form.

Condition (2) can take one of two forms. In one version, the students are given the rule: this has been termed "deductive" grammar learning. In another version, the students are asked to try to work out the rule on their own ("inductive" rule learning). The latter version is termed "rule search" in some of the studies discussed here.

I first review a set of experiments in which the impact of direct instruction in grammar is measured directly. This is followed by a discussion of individual studies focusing on the issues of consciousness raising and whether formally learned competence becomes "automatic." The final study reviewed here confirms the limits of "metalinguistic" knowledge.

Stephen D. Krashen (Ph.D., UCLA) is Professor of Education, University of Southern California, Los Angeles.

Experimental Studies Claiming to Show an Effect for Grammar

These studies have several characteristics in common:

1. Subjects were experienced adult language students, which means they were used to direct teaching of grammar, expected it, and had survived it.

2. Comparison groups had either no treatment at all, or received what can only be described as impoverished comprehensible input. Subjects, we are told, were focused on meaning, but it was always in an extremely contrived situation, in which context and interest was minimal. In addition, it is clear in some of the studies that some of the students in this condition were focused on form. The studies, thus, investigated only the impact of more direct instruction (more focus on form and more explicit presentation of rules) versus less, not "learning" versus "acquisition," contrary to what some of the investigators claim.

3. Tests were given soon after the treatment was finished; thus, only the short-term effect of conscious learning was investigated.

4. On all measures used, the focus was on form, or correctness. In addition, in all measures, subjects had time to apply the rules they

had studied. While it is claimed that the "time" condition was not satisfied in several of the studies, in only one case was time pressure quantified, and I argue that these results show that there was sufficient time for rules to be applied.

The consistent result is that those who had more rule-based instruction and form-focus did better, but in nearly all cases the effect is quite modest and some cases it is completely absent. The studies thus only show that more instruction means a bit more consciously learned competence, a conclusion that is consistent with the claims of the Monitor hypothesis (Krashen 1982).

Master (1994): Subjects were university ESL students at UCLA and California State University, Fresno. All were considered to be at the intermediate level and most had studied the target rule, the English article, before.

Experimental subjects received six hours of systematic direct instruction on the article over a nine-week period. The same measure was used as a pre- and posttest, a fill-in-the-blank test in which students supplied the correct form of the article. Master provides these examples, some involving just one sentence:

> Carlos is ___ student at our university.

some involving pairs of sentences;

> Once there were many trees here. Now, ____ trees are gone.

and others involving a paragraph;

> ____ favorite food of ___ jaguar is ____ wild pig. ____ wild pigs move in ___ bands of fifteen to twenty. They have ____ great courage and ____ strength in ____ groups.

The measure clearly focuses students on form, especially when one considers that the experimental students had just had six hours of intensive work on the article system before taking the posttest. Master suggests, however, that the test measured acquisition, because "subjects were given the test without prior announcement and they were only given enough time to answer without deliberating upon their responses. It was hoped that the test would thus reflect spontaneous knowledge" (232). The nature of the time constraint was not discussed in any more detail.

Table 1 (see page 255) provides the results of the original study (UCLA) and the replication (Fresno). Master reports that experimental group gains were statistically significant, but comparison group gains fell just short of significance, reaching the one-tail .10 level.

Most important, the gains were very modest. After six hours of intensive study, the two experimental groups gained only 6.5 percent and 9 percent. Master also calculated the effect size for the difference in gain scores between the experimental and control groups in the UCLA study, based on the means and standard deviations, and reported an effect size (d) of .664.

Rather than demonstrating that instruction works, Master has confirmed the limits of conscious learning: Using subjects who are supportive of and experienced with grammar learning (international students at the university level), and who underwent intensive study of the target rules, and using a discrete-point grammar test focused exclusively on the target rule administered very soon after the treatment, gains were very modest. The average UCLA student went from about a "C" grade to a low "B," while Fresno students went from about a "C-" or "D+" to a "C."

Master also reported no significant correlation between amount of formal study and performance on the pretest, confirming the inefficacy of instruction. He also reported no significant correlation between length of residence in the United States and article pretest performance. It would be interesting to see if amount of pleasure reading and article use are related, as other studies have reported a clear relationship between acquisition of complex syntactic forms and reading (Lee, Krashen, and Gibbons 1996; Stokes, Krashen and Kartchner 1998).

Leeman, Aregagoitia, Fridman and

Doughty (1995) examined the impact of focusing on form without explicit rule teaching. They asked sixth-semester university level Spanish students in a "focus on form" condition to read passages with target verb forms (preterit and imperfect) underlined and highlighted, with different colors for different forms, while paying special attention to verb forms and their meanings. Students then answered questions based on the passage while paying attention to form, then discussed the readings and questions in class, paying attention to the formal aspects of their output and to the teacher's corrections, then participated in a debate while focusing on correctness of verb forms, and finally viewed the debate on video and evaluated their classmates' performance. All students had studied these verb forms previously.

Their performance on the target forms was then compared with students who had undergone similar activities without a focus on form. Results are presented in Table 2 (see page 255): There was no difference at all in gains between pre- and posttests for an essay (pre- given one week before the treatment; post- given five weeks afterwards) and on a cloze test that focused on the target items: Neither group showed significant gains. The focus-on-form group showed a significant gain for accuracy in using the target verb form in the debate (predebate performed two weeks before; postdebate performed one week afterwards).

Even this gain, however, does not demonstrate acquisition. First, only five subjects did both the pre- and postdebate. Three of the subjects hardly produced any verb forms with the imperfect in either debate, and three produced fewer than ten instances of the preterit on the postdebate. Thus, Leeman et al.'s results are really based on the performance of at most three subjects.

One can argue that the time constraints of the debate prevented the subjects from applying their conscious knowledge. But by the time subjects did the final (third) debate, they had done two debates before, and had discussed and read about the material. They were thus prepared to at least some extent. In addition, we do not know how great the time pressure was during the actual performance of the debate.

At best, this study shows only that one can increase accuracy in semiprepared oral presentations for a few college students who were survivors in Spanish (sixth semester) after a strong dose of focus on form on a rule they have already studied. The finding that there was absolutely no gain for the other two measures is strong evidence that the treatment was not effective.

Robinson's (1995) subjects were university students enrolled in intermediate ESL classes, which strongly suggests that they were accustomed to formal grammar learning and expected it. Robinson employed four conditions, but the results of only three are discussed here (in one condition, "implicit learning," subjects viewed sentences and were then tested on the position of words appearing in the sentences).

In all conditions, subjects viewed twenty sentences exemplifying a "hard" rule of English (pseudo-cleft) and ten exemplifying an "easy" rule (subject-verb inversion is allowed with adverbials of location, but not adverbials of time).

The three conditions of interest were these:

1. The incidental condition was considered to be "acquisition." After each sentence, subjects were asked to answer a yes/no question, intended to focus the subject on meaning. It must be pointed out, however, that this is an extremely impoverished acquisition environment, with practically no context and no message of interest. In addition, the experimental situation itself promoted a focus on form, as we will see below, and we have no idea whether the subjects were developmentally ready to acquire the target rules, that is, whether the rules were at "i+1."

2. In the rule-search condition, subjects were asked to try to figure out the rules. After each stimulus sentence, subjects were asked if they had made progress in doing so.

3. In the instructed condition, subjects read

explanations of the hard and easy rules. Questions after each sentence dealt with form, e.g., "Did the subject of the sentence come before the verb?"

Thus, each condition promoted successively greater focus on form, and provided greater knowledge of the target rule. If instruction is successful, its effects should increase with each condition, moving from (1) to (3).

The measure was a grammaticality judgment test given immediately after the treatment, with twenty sentences for each rule. Robinson does not present the raw data—my Table 3 (see page 255) was read from his Figure 1 (320). It was not possible to calculate effect sizes, because standard deviations were not provided, nor were precise p-levels reported for post-hoc comparisons.

While instructed learners were significantly better than the other groups for easy rules, they were only significantly better than rule-searchers for hard rules. Rule searchers were not better than incidental students.

As usual, the effect was modest: the instructed group got 17 right out of 20 on the easy rule, while the incidental group got 14.6 right out of 20. Also, the effect was only demonstrated to be short term, and was only significant for the "easy" rule, the one that was more consciously learnable.

Robinson reported that most subjects in all conditions were, in fact, focused on form: 15 of the 26 "incidental" subjects said that they looked for rules (compared to 20 in the "instruction" and "rule search" conditions; this difference was not statistically significant, chi square = .428 for df = 1; chi square = .556, df = 2; Robinson reported statistical significance for this comparison but included the implicit group as well). For all conditions, those who said they could verbalize the rule did a little better, but differences were not statistically significant. In other words, those who felt they knew the rule well did not do much better than those with a vaguer understanding.

Robinson thus does not show that learning is superior to acquisition. Rather, this study has little to do with acquisition and confirms only that experimental situations that focus students on form heavily and that provide explicit knowledge of the rule produce small advantages for accuracy on form-based tests in the short run. This study only confirms that we can make small improvements in processes that the brain does poorly in the first place.

Robinson (1997) used a similar design and has similar problems. Once again, subjects were intermediate level students of English as a second language at the university level, i.e., experienced formal learners. This time the target rule was dative alternation, i.e., with one syllable verbs one can say both:

(1) John gave the cake to Mary; and
(2) John gave Mary the cake.

But with verbs with more than one syllable, both versions are not possible:

(3) John donated money to the church.
(4) John donated the church money.

(Robinson notes that this is a simplified version of the actual rule.)

All subjects viewed 55 sentences, corresponding to sentences (1), (2) and (3) above, and nonsense words were used for verbs to control for prior knowledge. The incidental and instructed conditions were similar to those used in Robinson (1995). The third condition of interest to us here was an "enhanced" condition in which crucial aspects of the stimulus sentences were put in boxes to make them more salient: This condition thus encouraged focus on form without rule presentation. Table 5 (see page 256) presents accuracy results for novel grammatical and ungrammatical sentences on a grammaticality judgment task.

Similar results were found for accuracy and reaction time: The incidental and enhanced groups took about twice as long to make judgments. While the slight superiority for the instructed group in judging grammatical sentences is typical of what one finds in these studies, for ungrammatical sentences the advantage is much larger than usual.

Again, it is reasonable to assume that little acquisition took place in any of the conditions in this study. First, dative alternation is, most likely, a late acquired rule. Second, subjects in the incidental and enhanced conditions were given emaciated comprehensible input. Third, subjects in the incidental condition were also concerned with form: Robinson reported that about one-half of the subjects in the incidental and enhanced conditions said they tried to figure out the rule (241).

Clearly, the subjects in the instructed condition were the only ones who had extensive conscious knowledge of the target rule: No subject in the other conditions could verbalize the "critical factor" governing dative alternation (244). Compare this to Robinson (1995), in which many subjects succeeded in figuring out the rule (7/20 in the incidental and rule-search conditions). The small amount of acquisition and learning developed by the incidental and enhanced groups (as well as their previous knowledge) was apparently enough to confirm that sentences were grammatical, but this knowledge was not enough to make accurate judgments on ungrammatical sentences.

Once again, this study does not compare acquisition and learning. Rather, it compares the impact of different amounts of form-focus and consciously learned competence on performance on a grammaticality judgment test, given immediately after the treatment took place. The instructed group had had the most focus on form and knew the rule far better than the others. It is not at all surprising that they outperformed the other groups to such a degree.

In de Graaff's (1997) studies, subjects were also experienced formal learners: All were university students. They were described as monolingual speakers of Dutch, but all had studied English, French and German for four to six years in school.

Also, de Graaff describes the two conditions used in his study as similar to the rule-search and instructed conditions used in Robinson (1995). Two groups of subjects studied four target structures in an artificial language based on Esperanto for a total of fifteen hours. The "implicit" group was similar to Robinson's rule-search group. They were focused on form but were not provided with an explanation of the rules, while the explicit group was. Both groups received "immediate feedback" on the correctness of their responses during exercises. As in other studies of this kind, all measures focused the subjects on form. The grammaticality judgment test was, however, given under two conditions, one with time pressure: "participants were instructed to carry out the task as quickly as possible"(259).

I present de Graaff's results in Table 6 (see page 256). Once again, the data were read from graphs.

As indicated in Table 6, the explicit group was better on all measures, confirming that for form-based language learning, subjects who are given more information about form do better than subjects given less information about form. Considering that the treatment lasted fifteen hours and four structures were taught, differences were modest. There was little deterioration of performance five weeks later, which is also consistent with the results of other studies: It usually takes somewhat longer for loss to occur (Krashen 1994; see discussion below).

In de Graaff's study it was also reported that subjects did slightly better on the grammaticality judgment task when they had no time pressure, but the difference was statistically significant for only one of four target structures. Fortunately, de Graaff measured reaction times in both conditions. With time pressure, subjects took an average of 7 seconds on the immediate posttest and 6.3 seconds on the delayed posttest. Without time pressure, they took 12.7 and 8.9 seconds. Thus, both time pressured and nontime pressured conditions may have allowed enough time to access the conscious Monitor, as de Graaff notes (271). The time pressure condition may not have provided as much pressure as real conversation, in which the speaker must contend with the temporal demands of

having a conversational partner.

The results of de Graaff's study are thus fully consistent with the supposition that only learning was involved in this study.

The most recent attempt to show that direct instruction works is Manley and Calk (1997). Thirteen university students enrolled in third-year French participated in four different lessons, each focused on a different target rule. Rules used were those found to give students problems on an essay done before the treatment. Manley and Calk attempted to base each lesson on a different philosophy of instruction, but the similarities far outweighed the differences: In all cases, there was an explicit presentation of the rule and practice using the rule in output activities, with a clear focus on form.

Table 7 (see page 256) presents the number of errors students made on a composition written just before each lesson and on a composition written at the end of the semester.

On the basis of four separate chi square analyses, Manley and Calk conclude that grammar study helped in three out of four cases. Raters did not, however, consider the final composition to be of better overall quality than the first composition students wrote.

Once again, subjects were experienced rule learners. Nine of the thirteen felt that the grammar presented in class was useful, with four indicating it was only "somewhat" useful and none saying it was useless. In addition, subjects had been focused on form during the treatments, and knew they were evaluated on accuracy.

A very serious flaw in this study, however, is that we do not know the length of the compositions and how many times the target structures were attempted. Also, the presentation method used in Manley and Calk's table gives the impression that the impact of instruction was much larger than it was, because only the total number of errors was given. In Table 7, I also include the average number of errors per subject, which is the usual way this kind of data are reported. The impact of instruction appears very tiny when presented this way: For noun/adjective agreement, average errors per subject fell from about 3 to about 1, for possessive adjectives, from about 1 to 1/3, and for the definite article, from about 1 to about 1/2. We thus see very few errors overall, and a very small effect of instruction.

Summary of Studies

As noted at the beginning of this paper, all subjects in all of the studies discussed here were experienced "learners": intermediate ESL students at the university level (Master 1994; Robinson 1995, 1997), intermediate or advanced foreign language students at the university level (Leeman et al. 1995; Manley and Calk 1997), and university students with a substantial amount of experience studying foreign languages (de Graaff 1995).

In some cases, comparison groups experienced what is labeled "implicit learning," but it is not correct to describe these conditions as acquisition-rich (Robinson 1995, 1997; de Graaff 1995): Only isolated sentences were presented, and subjects were quizzed on their content. There is also evidence that because of the contrived, artificial situations, many subjects in the implicit condition were focused on form, and in one study a substantial number of "implicit" subjects said they could state the rule (Robinson 1995).

Subjects were focused on form on all tests. Master (1994) used a fill-in-the-blank grammar test; Robinson (1995, 1997) and de Graaff (1995) used grammaticality judgment tests. Leeman et al. used three measures and it can be argued that there was considerable form-focus on all three: the "cloze" was actually a fill-in-the-blank test specifically focused on the target forms, and students in the form-focus condition had recently done debate and essay activities in the treatment with a focus on getting the target items correct. Since they were also used as tests, it is reasonable to hypothesize that students knew that accuracy in the use of the target forms was the name of the game. Manley and Calk (1997) also used an essay for pre- and posttesting, but it is very likely that students realized that form was the issue.

In three instances, there was an attempt to

induce some time pressure, but in the one case in which this was quantified (de Graaff 1995), it appears that there was enough time for the application of consciously learned rules.

These studies, thus, do not compare direct instruction and comprehensible input. Rather, they compare the impact of more or less direct instruction on tests designed to measure conscious learning. Predictably, they show that more focus on form and more information presented about rules results in more conscious learning.

But not much. One can quantify the effect of conscious learning in several ways, which I have attempted to do in Table 8 (see page 257). (It was not possible to do a true meta-analysis involving effect sizes, because, as noted above, it was not possible to do the calculations for several studies.) Percentage gain of posttest over pretest may be the easiest to interpret, as it translates readily into classroom practice. As seen in Table 8, subjects in the studies discussed here show only a modest improvement with more direct instruction, when pre- and posttests can be compared. The only impressive gain is the debate in Leeman et al., but as noted earlier, only three subjects contributed meaningful data for this comparison. When we examine the advantage calculated by comparing the posttest performance of experimental and comparison groups, the results are similar, with only one condition in Robinson (1997) showing impressive results; judgment of ungrammatical sentences. While some researchers are apparently impressed when their subjects show a 6.5 percent gain after six hours of treatment on a set of rules, few teachers would be satisfied with such results.

Thus, while investigators in all of these studies claimed that conscious learning triumphed, no data are provided that hint that conscious learning has more than a peripheral effect; nothing has changed in the theory of language acquisition.

Consciously Learned Competence Fades

As noted earlier, in the studies discussed in this section, testing was done immediately after the treatment ended. The exception is de Graaff (1995), who also included a delayed posttest five weeks later. A delay of this length may not be long enough to see fading of consciously learned competence. The time it takes for learned knowledge to disappear appears, however, to be a function of the intensity of training. (Manley and Calk's posttest could also be considered delayed, but we are not informed how much time there was between the treatment and the final essay. Interestingly, the structure in which accuracy decreased was the one taught first, with the longest delay between treatment and test.)

Scott and Randall (1992) reported a drop-off in accuracy on delayed tests administered only four weeks after the end of the treatment. In their study, first-year French students studied three rules and the treatment was short: "The grammar lessons included two pre-reading exercises, an introductory dialogue illustrating the meaning of the targeted grammar structure and a one-sentence rule followed by examples in context with translations" (358).The immediate and delayed tests contained multiple-choice and completion exercises, as well as a task in which students had to write "personalized" sentences using the structure (359).

As in previous studies, subjects showed clear drops in accuracy on the second test. In this study, however, the decline appeared more rapid than the decline seen in other studies; this may be due to the fact that the treatment was very short, only about four minutes per rule (359).

Working much harder, however, only delays the inevitable: While Day and Shapson's subjects had six weeks of instruction on the French conditional and held their gains for eleven weeks (Day and Shapson 1991), Harley's subjects spent eight weeks (about twelve class hours total) on the passé composé and imparfait, but they lost their advantage over a comparison group on tests administered three months later (Harley 1989). Subjects studied in White (1991) had five weeks of instruction on adverb placement

and held their gains for five weeks, but had lost them when tested one year later. An apparent exception is Spada and Lightbown (1993), whose subjects had nine hours of instruction on English question formation over two weeks, and actually showed some improvement on a posttest administered six months after the instruction. The comparison group also improved at a comparable rate on the target structure during this time, however.[1]

Consciousness Raising?

Fotos (1993) investigated the role of grammar study in "consciousness raising." Her subjects, 160 EFL students in Japan, were divided into three groups: One group did grammar tasks in which there was a focus on grammatical form, a second group had traditional grammar lessons, and a third group participated in communicative tasks in which there was no focus on form. After each treatment, one for each target rule in groups 1 and 2, all subjects were asked to do a "noticing task" in which they read a story and were asked to underline any "special use" of English. One week later, they did a similar noticing task with dictation. The story and dictation contained exemplars of the target structures used in the grammar task and grammar lesson. After three weeks, subjects in group 1 and 2 took a grammaticality judgment test and production tests that focused on the target structures (unscramble sentences, sentence combining).

Fotos reported that subjects in the first two groups were better able to notice examples of the target structures in the noticing task. She reported, however, no relationship between the ability to notice and combined scores on the proficiency measures (with the exception of the grammar lesson group, and only for one structure out of three, indirect object placement: The correlation was modest, $r = .35$.). In addition, there was clearly less noticing on the second administration of the noticing task, one week after instruction. Moreover, noticing frequency was not high, with subjects from the focus on form and grammar groups noticing about two to three items out of five. Fotos' results, in my view, provide good evidence that consciousness raising does not play a role in language acquisition. (For the record, there was no difference between the grammar task and grammar lesson groups on the grammar test; the communicative task group did not take the grammar test.)

Does Learning Become Acquisition?

DeKeyser (1997) very clearly deals with conscious learning and not acquisition. Subjects studied an artificial language and were supplied with "a traditional presentation of grammar rules" (201). In addition, care was taken to insure that subjects understood the rules. This was followed by practice activities in which the focus was on correctness at all times, with constant feedback: "each error triggered an automatic explicit explanation of what grammatical categories in the answer were wrong and why ..." (203).

DeKeyser found that learning curves for accuracy and reaction time showed a sudden improvement at the beginning, and then a number of slower improvements, consistent with the hypothesis that knowledge in early stages was "declarative" and then was gradually automatized: Using different terminology, the claim is that "learned" competence becomes "acquired" competence. In addition, the learning curve was nearly identical to that found for learning other cognitive skills, such as algebra.

DeKeyser concludes from this finding that "the learning of second language grammar can proceed very much in the same way that learning in other cognitive domains, from geometry to computer programming, has been shown to take place" (214), and that second-language acquisition is not "necessarily acquired through the implicit mechanisms of a separate mental module (as is generally accepted for first-language acquisition)" (213).

But DeKeyser's results can also be interpreted as showing that conscious learning and subconscious acquisition of language are different. I have suggested (Krashen 1985) that conscious learning is done outside the "language acquisition device" and utilizes mechanisms used in other areas of cognition.

DeKeyser's results are consistent with this suggestion: What is obviously language learning, not acquisition, showed a pattern similar to that seen in other areas of cognition.

DeKeyser also found that practice in comprehension tasks did not transfer well to production, and vice versa. He concluded that this result "is contrary to the idea of linguistic competence acquired through comprehension being equally available for production and comprehension" (213). In my view, these findings only confirm that in this study we are dealing with learning, not acquisition.

More Evidence for the Limits of Conscious Learning

Alderson, Clapham and Steel (1997) administered a variety of tests to 509 first-year students of French at the university level and found two factors, "the first one loading on metalinguistic and the second on...language proficiency variables" (115).

The "metalinguistic measures" included the following:

1. Parts of speech: read a sentence and identify the parts of speech, given in both French and English.
2. Given an ungrammatical sentence, correct the sentence and state the rule being broken. For the French sentences, one point was given if the error was identified and corrected (labeled "French errors"), one point if "metalanguage" was used, and one point for a correct statement of the rule ("French rules"). For English sentences, one point was given for correction of the error, one for an accurate rule ("English rules"), and one "for giving the reason" ("English reasons").

Subjects also took a number of proficiency tests of French, including an 18-time rational cloze test ("cloze"), a grammar test, and a listening comprehension test.

"Metalinguistic knowledge" corresponds to conscious learning, and the clear two factor solution, presented in Table 10 (see page 257), confirms that this kind of knowledge is not strongly related to actual performance, even on tests in which the focus is on form.

Alderson et al. conclude that "there is no evidence from this study to justify the teaching of metalinguistic knowledge as a means of improving students' linguistic proficiency" (118).

Summary and Conclusion

This review is not exhaustive of all studies in which the claim is made that grammar study is good for students. I have attempted only to discuss some current examples (see also Krashen 1993, 1994). Nor have I discussed the evidence supporting the alternative, comprehensible input, as this has been done in many other places (e.g., Krashen 1985, 1994, 1997). My goal was only to illustrate that recent studies claiming to support grammar teaching over subconscious acquisition really show nothing of the sort.

The Monitor hypothesis (Krashen 1982) claims that several conditions are necessary for the successful application of consciously learned rules of grammar:

(1) Knowledge of the rule;
(2) Concern with correctness, or focus on form;
(3) Sufficient time.

What is shown here is that even when we optimize the conditions for Monitor use, that is, give students plenty of instruction on few target rules, and test them on measures in which they are focused on form and have time to apply the rules, we see little effect: The Monitor is weak (but not completely useless; see Krashen 1982).

The studies reviewed here are of value, but their focus, in Smith's terms (1996), is on what the brain does not do well.

NOTES

[1] The comparison group teacher promoted a focus on form, frequently correcting students' use of question forms, but students in the experimental classes produced more questions and had more feedback. Spada and Lightbown point out that the comparison teacher might have emphasized form more in the months preceding the treatment,

which in their view explains why this group also did well on the delayed posttest. Comparison students, however, also heard far more questions (Table 3, 214). Clearly, this one study does not help us decide among competing hypotheses.

REFERENCES

Alderson, J.C., C. Clapham, and D. Steel D., "Metalinguistic Knowledge, Language Aptitude and Language Proficiency." *Language Teaching Research* 1: 93-121.

Day, R. and S. Shapson. 1991. "Integrating Formal Approaches to Language Teaching in French Immersion: An Experimental Study." *Language Learning* 41: 25-58.

de Graaff, R. 1997. "The eXperanto Experiment: Effects of Explicit Instruction on Second Language Acquisition." *Studies in Second Language Acquisition* 19: 249-276.

DeKeyser, R. 1993. "The Effect of Error Correction on L2 Grammar Knowledge and Oral Proficiency." *Modern Language Journal* 77: 501-514.

DeKeyser, R. 1997. "Beyond Explicit Rule Learning: Automatizing Second Language Morphosyntax." *Studies in Second Language Acquisition* 19: 195-221.

Fotos, S. 1993. "Consciousness Raising and Noticing through Focus on Form: Grammar Task Performance Versus Formal Correctness." *Applied Linguistics* 14: 385-407.

Harley, B. 1989. "Functional Grammar in French Immersion: A Classroom Experiment." *Applied Linguistics* 10: 331-359.

Krashen, S. 1982. *Principles and Practice in Second Language Acquisition.* New York: Prentice Hall.

Krashen, S. 1985. *The Input Hypothesis.* Beverly Hills, CA: Laredo Publishing Company.

Krashen, S. 1992. "Under What Conditions, If Any, Should Formal Grammar Instruction Take Place?" *TESOL Quarterly* 26: 409-411.

Krashen, S. 1993. "The Effect of Formal Grammar Study: Still Peripheral." *TESOL Quarterly* 27: 722-725.

Krashen, S. 1994. "The Input Hypothesis and Its Rivals." 45-77 in N. Ellis (Ed.) *Implicit and Explicit Learning of Languages.* London: Academic Press.

Leeman, J., I. Aregagoitia, D. Fridman, and C. Doughty. 1995. "Integrating Attention to Form with Meaning: Focus on Form in Content-Based Spanish Instruction." 217-258 in R. Schmidt, ed., *Attention and Awareness in Foreign Language Learning.* Hawaii: University of Hawaii Press.

Light, R. and D. Pillemer. 1984: *Summing Up: The Science of Reviewing Research.* Cambridge: Harvard University Press.

Manley, J. and L. Calk. 1997. "Grammar Instruction for Writing Skills: Do Students Perceive Grammar As Useful?" *Foreign Language Annals* 30: 73-81.

Master, P. 1994. "The Effect of Systematic Instruction on Learning the English Article System." 229-252 in T. Odlin, ed., *Perspectives on Pedagogical Grammar.* Cambridge: Cambridge University Press.

Robinson, P. 1995. "Aptitude, Awareness, and the Fundamental Similarity of Implicit and Explicit Second Language Learning." 303-357 in R. Schmidt, ed., *Attention and Awareness in Foreign Language Learning.* Hawaii: University of Hawaii Press. 303-357.

Robinson, P. 1997. "Generalizability and Automaticity of Second Language Learning under Implicit, Incidental, Enhanced, and Instructed Conditions." *Studies in Second Language Acquisition* 19: 223-247.

Scott, V. and S. Randall. 1992. "Can Students Apply Grammar Rules for Reading Textbook Explanations?" *Foreign Language Annals* 25: 357-367.

Smith, F. 1996. *Between Hope and Havoc.* Portsmouth: Heinemann.

Spada, N. and P. Lightbown. 1993. "Instruction and the Development of Questions in L2 Classrooms." *Studies in Second Language Acquisition* 15: 205-224.

Stokes, J., S. Krashen, and J. Kartchner. 1998. "Factors in the Acquisition of the Present Subjunctive in Spanish: The Role of Reading and Study." ITL: Review of Applied Linguistics.121-122:19-25.

White, L. 1991. "Adverb Placement in Second Language Acquisition: Some Effects of Positive and Negative Evidence in the Classroom." *Second Language Research* 7: 133-161.