The Role of Positive and Negative Evidence in the Classroom: Does the Subset Principle Operate in L2 Learning?

Katsumasa Shimada

1. INTRODUCTION

1.1 Background

The Subset Principle is a learning principle which attempts to guarantee that first language (L1) acquisition can be achieved with positive evidence only. It claims that when the two grammars meet the Subset Condition, that is, when they are in a subset/superset relation, if input which could be accommodated by either of two grammars meeting the condition is given, the most restrictive grammar consistent with the input should be adopt. Actually, negative evidence is not available in L1 acquisition, and it does not play a significant role, if any.¹

The question arises as to whether the principle still operates in second language (L2) acquisition, especially in the input-poor environment like L2 learning in the classroom. One of the possible answers to the question is that the Subset Principle operates in L2 learning as it does in L1 acquisition, that is, L2 learners generate a subset grammar with positive evidence alone, which is referred to as the subset hypothesis.

In contrast to this view is a position which assumes that L2 learn-
ers no longer apply the Subset Principle, instead they will be influenced by their L1. This position is referred to as the transfer hypothesis. (White, 1989a; 1989b)²¹

Over the past few years, a number of studies have been conducted on which position is better to explain L2 learning. White (1989a) investigated adverb placement by French-speaking learners of English and the results supported the transfer hypothesis. White (1991) examined the effects of negative evidence in a classroom setting and argued that "negative evidence would be effective in helping them to arrive at the correct properties of English"(p.151). Trahey and White (1993) investigated the effects of an input flood of naturalistic positive evidence and showed that supplying positive evidence alone did not necessarily trigger the appropriate L2 value of a parameter.³⁰

1.2 Research questions

The three research questions are addressed in terms of the operationality of the Subset Principle in L2 learning in the classroom context. Against the results of Trahey and White's study described above, the question arises as to whether or not positive evidence alone is really insufficient to parameter-reset the value of L1. If ample input of positive evidence is intensively supplied, then what will happen in L2 grammar? In other words, does intensive supply of positive evidence help Japanese learners to parameter-reset from L1 value to L2 value?

The next question will be concerned with the role of negative evidence. If negative evidence is supplied in addition to positive evidence, what will happen in L2 grammar? Is it more effective to reorganize L2 grammar than positive evidence alone? More concrete-
The Role of Positive and Negative Evidence in the Classroom

ly speaking, do Japanese learners of English change from the superset grammar (L1) to the subset grammar (L2) more effectively with negative evidence than with positive evidence alone?

Schwartz (1993) claimed that L2 grammar reorganization should affect the other properties of clustering and that a set of deductive consequences should emerge automatically. Our concern will be the effects of negative evidence on what learners have not learned in the classroom, or what emerges automatically as a result of learning. If positive evidence alone is insufficient and negative evidence is indispensable for parameter resetting to take place, then in what way negative evidence is effective in helping learners to reset the L1 value of a parameter? We will examine the positive effects of negative evidence on the transfer of learning. Positive transfer of learning is defined as the effect of learning in one situation which helps or facilitates learning in another situation. If negative evidence is provided, do the learners learn what they have not been taught? Does exposure to negative evidence produce more effect of learning transfer than exposure to positive evidence alone?

To summarize, the purpose of this study is to investigate whether positive evidence alone can trigger parameter resetting, or negative evidence is indispensable to reset the parameter value once set in L1 acquisition and to seek what kind of teaching can help Japanese learners to parameter-reset effectively.

1.3 Hypotheses

When Japanese (L1) and English (L2) are in a superset/subset relation, it seems that Japanese learners of English cannot abandon the superset grammar and cannot create the subset grammar easily.
Let us examine the case of adverb placement.

(1) a. John often watches television. SAVO(+adj)
   b. *John watches often television. SVAO(−adj)

(2) a. John-wa yoku terebi-o miru. ŠAVO(+adj)
   b. John-wa terebi-o yoku miru. SVAO(−adj)

English adverbs cannot intervene between a verb and its direct object, as in (1b), while Japanese counterpart can, as in (2b). This is called adjacency condition on case assignment, which is a requirement that an NP receiving case must be next to its case assigner. English (1b) sentence is ungrammatical because it violates the adjacency condition, whereas Japanese (2b) sentence is grammatical even if the adverb intervenes between the verb and its object. The similarity and difference between the two languages illustrated (1) and (2) can be explained by the value of a parameter of Universal Grammar. Adopting Chomsky’s assumption that there is parametric variation, English is [+strict adjacency] and Japanese is [−strict adjacency] in terms of adverb placement.⁴ It is predicted that Japanese learners of English are likely to violate the adjacency condition based on their L1 knowledge. In other words, negative transfer from their native language will cause the violation of adjacency condition.

The following hypotheses will be investigated in this paper:

(1) Providing Japanese learners with input of positive evidence intensively would help them to parameter-reset from L1 value (−strict adjacency) to L2 value (+strict adjacency).

(2) Japanese learners of English would change from the superset grammar (L1) to the subset grammar (L2) more effectively with negative evidence than with positive evidence alone.

(3) Exposure to negative evidence would produce more learning trans-
The Role of Positive and Negative Evidence in the Classroom

er from adverb to prepositional phrase placement than exposure to positive evidence alone.

2. Method

2.1 Subjects

Seventy-four freshmen enrolled in English at Meijo University participated in this experiment. Subjects consisted of two groups: Class 1 and Class 2. Thirty-nine students (Class 1) majored in mathematics, and thirty-five (Class 2) were traffic mechanics majors. Since the research was conducted over a three-month period, learner absences were not avoidable. Consequently, if learners missed any one of the grammar lessons or the grammar task performances, their scores for the performance were not included in the data analysis. The final numbers of the subjects who received the grammar lesson and sat for all the tests were thirty for each of the two groups.

Although standardized test scores were not available, judging from the classroom observation and the results of the pre-test, these students were considered to be beginner-level or false beginners.

Before the research began, the two classes were administered a rational cloze test which included 30 deleted words, in order to investigate whether there were significant differences in English proficiency. Unpaired t-test was performed and did not indicate a significant difference between the test score means (t=0.163, p=0.865). Therefore, the two classes were considered to be identical before the treatment.

2.2 Procedures

At the beginning of the experiment, both groups were pre-tested
on adverb and prepositional phrase placement. Three different kinds of tests: (a) grammaticality judgment test, (b) comparison test, and (c) manipulation test were administered to measure their syntactic knowledge on the positions of adverb and prepositional phrase in a sentence.

The study involved two kinds of treatment as the experimental conditions. Two weeks after the pre-test, information gap activities were given to a group (Class 1), which intensively provided ample input of positive evidence.\(^3\) Another class (Class 2) included rule finding and error correction activities, which provided negative as well as positive evidence as input. Immediately after the 30-minute class, the same tests were given again as a post-test to measure the immediate, or short-term effect of the treatments. Nine weeks after each grammar treatment, the same tests were administered again as a final-test to see whether there was the long-term effect of the learning.

2.3 Teaching materials and methods

Teaching focused on the positions of two kinds of English adverb: adverbs of manner (e.g., slowly, quickly, quietly, carefully) and adverbs of frequency (e.g., sometimes, often, usually, always). It should be noted that neither the information gap nor the rule finding group did not deal with the positions of prepositional phrases.

The characteristics of Communicative Language Teaching is to reflect the features of the real communication processes. Morrow (1981) illustrates three features of communication processes: information-gap, choice, and feedback. On the basis of the principles of the communicative methodology, an information exchange task was
developed, the purpose of which was to exchange information to bridge the gap between a pair of participants. For the information gap group, half of the class was handed out the material for Student A (Sheet A) and the remaining half was given the material for Student B (Sheet B). Sheet A had information that Sheet B did not have, and vice versa. The object of the task is to exchange the information they have so that they may know whether or not the descriptions presented in their own sheet is correct. Students worked in pairs. They were required to keep their own sheet out of sight of other participants so that it could not be seen by them. During the task each student transformed declarative sentences into interrogative or negative sentences at least 11 times, without changing the word order of S, V, O, and A. Student A’s output provided Student B with positive evidence, and vice versa.

An example of the information exchange activities is as follows:
A: Did Peter open the door slowly?
B: No, he didn’t. He didn’t open the door slowly.

Ellis (1993) claims that grammar teaching should be directed at consciousness-raising rather than practice. He refers to consciousness-raising as “a deliberate attempt on the part of the teacher to make the learners aware of specific features of the L2” (p.109). Ellis (1992) states that “it contributes to the processes of noticing and comparing” (p.238).

On the basis of the Comparative Principle (Carter, 1994), a rule finding task and error correction task were devised for the other group. In the rule finding task, students were shown five sets of sentences. Each set had a pair of Japanese sentences, which were both correct, and a pair of English sentences, one of which was correct
and the other of which was incorrect, because of violation of the adjacency condition. An example of the sentence sets is as follows:

(3) a. Peter-wa yakkurito to-o aketa.
    b. Peter-wa to-o yakkurito aketa.
    c. Peter opened the door slowly.
    d. *Peter opened slowly the door.

They were required to compare the Japanese sentence with the English equivalent, to compare the correct English sentence with the incorrect sentence paying attention to the differences of the position of adverb, and to generate three rules explaining the possible positions of the English adverb. Then they corrected 17 ungrammatical English sentences. It took about 30 minutes to complete the consciousness raising tasks. The materials can be founded in the Appendix A.

2.4 Tests

As described above, three different kinds of test were devised to conduct this experiment. All of the tests tried to measure learners' syntactic knowledge on the position of adverbs or prepositional phrases in a sentence. The grammaticality judgment test and the comparison test tried to measure the grammatical competence on the level of comprehension, while the manipulation test was devised to try to measure the syntactic knowledge at the production level.

They involved not only ungrammatical SVAO sentences but also SVPPO sentences. The latter sentence pattern was shown in neither the information gap activities nor the rule finding and error correction tasks. This means that each of the tests was devised to try to measure not only the direct effect of learning but also indirect effect of learning or learning transfer. Transfer effect was considered to
The Role of Positive and Negative Evidence in the Classroom

be able to be tested by measuring whether learners can apply a rule into the other environment. In this study, it was tested whether the rule of adverb placement can be applied to the sentence involving a prepositional phrase. The rule is that adverbs cannot intervene between a verb and its direct object. There is an analogy that prepositional phrases cannot appear between a verb and its object, either, since they grammatically function in the same way as adverbs. What has to be noticed is that the rule on the positions of prepositional phrase was not dealt with in the teaching phase at all.

In all the tests simple vocabulary and sentence structures were used. Some samples of the tests can be found in the Appendix B.

2.4.1 Grammaticality judgment task

The grammaticality judgment task required subjects to judge whether or not the sentence presented was grammatically correct in a limited amount of times. They read the sentence to be judged, and at the same time, heard the researcher reading the sentence aloud.

The test involved randomized 24 sentences, 12 of which were grammatical sentences paired with the ungrammatical sentences violating the adjacency condition. They consisted of 6 SVAO sentences involving 2 adverbs of manner, frequency and time, respectively, and 6 SVPO sentences.

In scoring the responses of the test, a very simple procedure was used. One point was given for each correct answer and a score of zero was given each incorrect response.

2.4.2 Comparison task

Another type of grammaticality judgment task took the form of
a comparison test, which consisted of 15 pairs of sentences. Each pair had two options, which differ only in syntactic form, as in (5)
(5) a. Jane often eats ice cream.
   b. Jane eats often ice cream.
Eleven pairs out of 15 were test pairs, in which there were 6 ungrammatical SVAO and 4 SVPPO sentences violating the adjacency condition and an ungrammatical SVAPP sentence locating an adverb of frequency after an intransitive verb. There were 4 control pairs that involved comparing sentences with adverbs or prepositional phrases in different positions in the sentence. Neither of them violated the adjacency condition.

Subjects were presented pairs of sentences and asked to decide whether the first or second sentence of the pair seems correct, or whether they both seem correct. The comparison test had an advantage in that the subjects had only to focus on the syntactic difference of the two sentences presented. Scoring procedure was the same as that of the grammaticality judgment test.

2.4.3 Manipulation task

In the manipulation task, subjects were presented four scrambled sentences involving four words or phrases and required to rearrange them into correct order. Of the four sentences, one sentence did not contain an adverb, instead involved an prepositional phrase. The remaining three contained an adverb of either manner or frequency.

Three of the sentences contained a transitive verb with its direct object and one contained an intransitive verb with a prepositional phrase. It should be noted that the test involved a sentence pattern characterized as the dative alternation (see No. 3).
The Role of Positive and Negative Evidence in the Classroom

If the sentence completed was grammatically correct, one point was given. Partial scoring was not used.

Subjects took on average 10 minutes to complete each of the grammaticality judgment test and comparison test and the manipulation test took 5–10 minutes.

2.5 Data collection and analysis

The first hypothesis was tested in terms of differences among the mean scores of the pre-, post-, and final test which the information gap group took. In order to confirm or disconfirm the hypothesis, the mean scores that the information gap group gained at the three test phases were compared. A one tail paired t-test was administered to verify that there were significant differences between the mean scores of two tests pair out of the three: pre-test vs. post-test; post-test vs. final test; pre-test vs. final test.

The second hypothesis was examined on the basis of the differences between the mean scores of the information gap and rule finding group. A one tail unpaired t-test was used to ascertain whether the differences between the mean scores of the two groups were statistically significant after the treatments.

The third hypothesis, namely that exposure to negative evidence would produce more learning transfer from adverbs to prepositional phrases than exposure to positive evidence alone, was examined by comparing the mean scores of the two groups with respect to the sentences involving a prepositional phrase. A one tail unpaired t-test was used to examine whether there were significant differences between them.

The data were analyzed with the number right scores and error
3. Results

3.1 Grammaticality judgment test

Table 1 gives the mean number right scores at the pre-test, post-test, and final test phase for the two groups. In principle, 24 is the maximum score, since there were a total of 24 sentences involving adverb and prepositional phrase placement in this task.

At the pre-test, there was no difference of the mean number right scores between the two groups (unpaired t=0.026, p=0.978). They both go up at the post-test and then drop at the final test. No significant difference was found to exist between the means of the post-test of the information gap and rule finding group (unpaired t=1.031, p=0.307). No significant difference was not observed between the mean scores of the two groups at the final test, either (unpaired t=1.328, p=0.186).

Let us pay attention to the errors the subjects made. The error scores are classified into the SVAO errors and the SVPPO error scores in terms of what constituent comes in the wrong position.

The SVAO error score represents the number of times the subject judges the ungrammatical SVAO sentence to be right. Table 2 shows the mean scores of the SVAO errors.

The error scores drop drastically in the case of both the information gap group and rule finding group at the post-test, although they both rise up at the final test. Since significant difference of the error scores between the two groups was observed at the pre-test phase (unpaired t=2.872, p=0.006), it did not make sense to compare the
The Role of Positive and Negative Evidence in the Classroom

### Table 1: Grammaticality judgment test; mean number right scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p</th>
<th>post-test</th>
<th>diff</th>
<th>p</th>
<th>final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>MS</td>
<td>14.47</td>
<td>3.87</td>
<td>***</td>
<td>18.33</td>
<td>-1.27</td>
<td>***</td>
<td>17.07</td>
<td>2.60</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3.15</td>
<td></td>
<td></td>
<td>3.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2</td>
<td>MS</td>
<td>14.43</td>
<td>4.47</td>
<td>***</td>
<td>19.20</td>
<td>-1.07</td>
<td>*</td>
<td>18.13</td>
<td>3.70</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.82</td>
<td></td>
<td></td>
<td>3.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>0.03</td>
<td></td>
<td></td>
<td>-0.87</td>
<td></td>
<td></td>
<td>-1.07</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

### Table 2: Grammaticality judgment test; mean SVAO error scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p</th>
<th>post-test</th>
<th>diff</th>
<th>p</th>
<th>final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>MS</td>
<td>3.23</td>
<td>-1.80</td>
<td>***</td>
<td>1.43</td>
<td>0.97</td>
<td>***</td>
<td>2.40</td>
<td>-0.83</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.41</td>
<td></td>
<td></td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2</td>
<td>MS</td>
<td>2.30</td>
<td>-1.67</td>
<td>***</td>
<td>0.63</td>
<td>0.77</td>
<td>**</td>
<td>1.40</td>
<td>-0.90</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.32</td>
<td></td>
<td></td>
<td>1.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>0.93</td>
<td></td>
<td></td>
<td>0.80</td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

### Table 3: Grammaticality judgment test; mean SVPPO error scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p</th>
<th>post-test</th>
<th>diff</th>
<th>p</th>
<th>final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>MS</td>
<td>2.60</td>
<td>-0.07</td>
<td>n.s.</td>
<td>2.53</td>
<td>0.20</td>
<td>n.s.</td>
<td>2.73</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.02</td>
<td></td>
<td></td>
<td>1.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2</td>
<td>MS</td>
<td>3.13</td>
<td>-0.17</td>
<td>n.s.</td>
<td>2.97</td>
<td>-0.27</td>
<td>n.s.</td>
<td>2.70</td>
<td>-0.43</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.18</td>
<td></td>
<td></td>
<td>1.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td></td>
<td>-0.53</td>
<td></td>
<td></td>
<td>-0.43</td>
<td></td>
<td></td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Mean scores between the two groups at the post-test as well as the final test phase.

Just like the SVAO error score, the SVPPO error score repres-
ents the number of times the subjects judge the ungrammatical SVPPO sentence to be right. The SVPPO error scores are presented in Table 3.

On pre-testing, both groups make statistically almost the same amount of the SVPPO errors (unpaired t=1.830, p=0.069). The SVPPO error score does not change in the case of not only the information gap but also the rule finding group at the post-test (paired t=0.220, p>0.05; paired t=0.492, p>0.05) as well as the final test (paired t=-0.943, p>0.05, paired t=0.774, p>0.05) phase.

3.2 Comparison test

Table 4 shows the mean scores of the comparison test at the three test phases. The maximum score on this task was 15, because there were 15 pairs on the comparison test.

At the pre-test, there is not significant difference of the mean scores between the information gap and rule finding group (unpaired t=0.142, p=0.883). The mean scores of both groups rise up at the post-test and fall down at the final test. The rule finding group performed better than the information gap group not only at the pre-test (unpaired t=4.410, p=0.000) but also at the final test (unpaired t=2.128, p=0.035).

The SVAO error score represents the number of times the subjects select the sentence containing the adverb in SVAO position or select a “both right” response in test pairs. Table 5 gives the mean SVAO error scores.

At the pre-test, there is not significant difference of the mean SVAO error scores between the information gap and rule finding group (unpaired t=1.115, p=0.269). The error scores drop dramatically in
The Role of Positive and Negative Evidence in the Classroom

Table 4: Comparison test; mean number right scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p post-test</th>
<th>diff</th>
<th>p final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 MS</td>
<td>7.70</td>
<td>1.97</td>
<td>** 9.67</td>
<td>-1.27</td>
<td>* 8.40</td>
<td>0.70</td>
<td>n.s.</td>
</tr>
<tr>
<td>SD</td>
<td>2.98</td>
<td></td>
<td>3.51</td>
<td></td>
<td>3.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2 MS</td>
<td>7.60</td>
<td>5.30</td>
<td>*** 12.90</td>
<td>-2.43</td>
<td>*** 10.47</td>
<td>2.87</td>
<td>***</td>
</tr>
<tr>
<td>SD</td>
<td>2.36</td>
<td></td>
<td>1.80</td>
<td></td>
<td>3.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td>0.10</td>
<td></td>
<td>-3.23</td>
<td></td>
<td>-2.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>n.s.</td>
<td></td>
<td>***</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 5: Comparison test; mean SVAO error scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p post-test</th>
<th>diff</th>
<th>p final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 MS</td>
<td>2.07</td>
<td>-0.73</td>
<td>** 1.33</td>
<td>0.23</td>
<td>n.s.</td>
<td>1.57</td>
<td>-0.50</td>
</tr>
<tr>
<td>SD</td>
<td>1.03</td>
<td></td>
<td>1.32</td>
<td></td>
<td>1.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2 MS</td>
<td>1.77</td>
<td>-1.60</td>
<td>*** 0.17</td>
<td>0.77</td>
<td>*** 0.93</td>
<td>-0.83</td>
<td>***</td>
</tr>
<tr>
<td>SD</td>
<td>1.02</td>
<td></td>
<td>0.52</td>
<td></td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td>0.30</td>
<td></td>
<td>1.17</td>
<td></td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>n.s.</td>
<td></td>
<td>***</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 6: Comparison test; mean SVPPO error scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p post-test</th>
<th>diff</th>
<th>p final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 MS</td>
<td>1.87</td>
<td>-0.50</td>
<td>n.s.</td>
<td>1.37</td>
<td>n.s.</td>
<td>1.40</td>
<td>-0.47</td>
</tr>
<tr>
<td>SD</td>
<td>1.26</td>
<td></td>
<td>1.43</td>
<td></td>
<td>1.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class 2 MS</td>
<td>1.77</td>
<td>-1.20</td>
<td>*** 0.57</td>
<td>0.73</td>
<td>*** 1.30</td>
<td>-0.47</td>
<td>*</td>
</tr>
<tr>
<td>SD</td>
<td>1.12</td>
<td></td>
<td>0.80</td>
<td></td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td>0.10</td>
<td></td>
<td>0.80</td>
<td></td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>n.s.</td>
<td></td>
<td>*</td>
<td></td>
<td>n.s.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

the case of both the information gap and rule finding group at the post-test (paired t=3.040, p<0.01; paired t=7.970, p<0.01), although they both rise up at the final test (paired t= -1.104, p>0.05; paired
t = -3.989, p < 0.001). The error score of the rule finding declines more drastically than the information gap group so that significant difference of the error score between the two groups was observed not only at the post-test (unpaired t = 4.403, p = 0.000) but also at the final test (unpaired t = 2.202, p = 0.045).

Just like the SVAO error score, the SVPPO error score represents the number of times the subjects select the sentence involving the prepositional phrase in SVPPO position or select a "both right" in the test pairs. The SVPPO error scores are presented in Table 6.

On pre-testing, both groups make almost the same amount of SVPPO errors (unpaired t = 0.319, p = 0.749). The SVPPO error score drops drastically in the case of the rule finding group at the post-test (paired t = 5.279, p < 0.001) although it rises up on the final test (paired t = -3.824, p < 0.001). In contrast, the information gap group's mean SVPPO error score remains statistically unchanged from the pre-test to the post-test (paired t = 1.630, p > 0.05).

3.3 Manipulation test

Table 7 shows the mean number right scores in the manipulation test. The maximum possible score on this task was 4 since each subject was given 4 scrambled sentences to be rearranged.

At the pre-test, there is no significant difference of the mean number right scores between the two groups (unpaired t = 1.803, p = 0.073). They both rise up at the post-test (paired t = -4.049, p < 0.001; paired t = -2.384) and remain unchanged at the final test (paired Table t = 1.400, p < 0.05; paired t = 1.532, p < 0.05). No significant difference was found to exist between the means of the post-test of the information gap and rule finding group (unpaired t =
The Role of Positive and Negative Evidence in the Classroom

Table 7: Manipulation test; mean number right scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p post-test</th>
<th>diff</th>
<th>p final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 MS</td>
<td>2.67</td>
<td>0.63</td>
<td>***</td>
<td>3.30</td>
<td>-0.23</td>
<td>n.s.</td>
<td>3.07</td>
</tr>
<tr>
<td>SD</td>
<td>0.98</td>
<td></td>
<td></td>
<td>0.78</td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Class 2 MS</td>
<td>3.10</td>
<td>0.33</td>
<td>*</td>
<td>3.43</td>
<td>-0.20</td>
<td>n.s.</td>
<td>3.23</td>
</tr>
<tr>
<td>SD</td>
<td>0.83</td>
<td></td>
<td></td>
<td>0.50</td>
<td></td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td>0.43</td>
<td></td>
<td></td>
<td>0.13</td>
<td></td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 8: Manipulation test; mean SVAO error scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p post-test</th>
<th>diff</th>
<th>p final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 MS</td>
<td>0.27</td>
<td>-0.27</td>
<td>***</td>
<td>0.00</td>
<td>0.13</td>
<td>***</td>
<td>0.13</td>
</tr>
<tr>
<td>SD</td>
<td>0.44</td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Class 2 MS</td>
<td>0.30</td>
<td>-0.30</td>
<td>***</td>
<td>0.00</td>
<td>0.13</td>
<td>***</td>
<td>0.13</td>
</tr>
<tr>
<td>SD</td>
<td>0.46</td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td>0.03</td>
<td></td>
<td></td>
<td>0.00</td>
<td></td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 9: Manipulation test; mean SVPO error scores

<table>
<thead>
<tr>
<th></th>
<th>pre-test</th>
<th>diff</th>
<th>p post-test</th>
<th>diff</th>
<th>p final test</th>
<th>diff</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 MS</td>
<td>0.57</td>
<td>-0.03</td>
<td>n.s.</td>
<td>0.53</td>
<td>0.00</td>
<td>n.s.</td>
<td>0.53</td>
</tr>
<tr>
<td>SD</td>
<td>0.50</td>
<td></td>
<td></td>
<td>0.50</td>
<td></td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Class 2 MS</td>
<td>0.47</td>
<td>0.10</td>
<td>n.s.</td>
<td>0.57</td>
<td>-0.07</td>
<td>n.s.</td>
<td>0.50</td>
</tr>
<tr>
<td>SD</td>
<td>0.50</td>
<td></td>
<td></td>
<td>0.50</td>
<td></td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>diff</td>
<td>-0.10</td>
<td></td>
<td></td>
<td>0.03</td>
<td></td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>n.s.</td>
<td></td>
<td></td>
<td>n.s.</td>
<td></td>
<td>n.s.</td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001

0.556, p=0.540). No significant difference was not observed between the mean scores of the two groups at the final test, either (unpaired t=0.686, p=0.502).
Let us examine the SVAO and SVPPO errors in the manipulation test. The sentences the subjects made were classified into 4 types: Type A–D, with respect to the position of adverb or prepositional phrase in the sentence. Type C errors come from the violation of the adjacency condition. Since there were only 4 question items in this test, we will see the SVAO errors in Question No. 1, and the SVPPO errors in Question No. 3.

Table 8 shows the mean scores of the Type C (SVAO) errors in Question No. 1, while Table 9 shows the mean scores of the Type C (SVPPO) errors in Question No. 3, which is referred to as the dative alternation sentence. The SVAO error scores drop drastically to zero at the post-test, and go up again at the final test, which is common for the two groups. It should be noted, however, that the test results for the SVPPO errors show no change for the two groups. The SVPPO error scores remain at about 0.5 through the three test phases.

4. DISCUSSION

The different task type, score type and test phase gives the different results for each hypothesis. The data we have examined indicate that learners' performance varies with the elicitation technique and learning stage, which makes it difficult to give a clear confirmation or rejection to the hypotheses.

Returning to the hypotheses investigated in this study, the first hypothesis, namely that providing Japanese learners with input of positive evidence intensively would help them to parameter-reset from L1 to L2 value in terms of strict adjacency, was partially supported,
The Role of Positive and Negative Evidence in the Classroom

since the results show that not only the rule finding but also the
information gap group performed better in the post-test than in the
pre-test, which was common in the three kinds of test. However,
the learning effect declined at the final stage, which was seen in the
two groups. The instruction to provide positive evidence does not
seem to have lasting effect on their L2 reorganization.

The results might be interpreted as the surface learning of the
sentence pattern. In other words, they might only have learned the
word order pattern at the surface level but might have not acquired
it. The results could be interpreted in another way, that is, the
short-term parameter resetting might be due to indirect negative evi-
dence with which learners would notice the non-occurrence of
ungrammatical sentences and conclude that they are not allowed.

The second hypothesis was that Japanese learners of English
would change from the superset grammar (L1) to the subset grammar
(L2) more effectively with negative evidence than with positive evid-
ence alone. The results in the comparison test partially support the
hypothesis.

Regarding the third hypothesis, namely that exposure to negative
evidence would produce more effect of learning transfer from adverbs
to prepositional phrases than exposure to positive evidence alone, the
results in the comparison test support the hypothesis, at least, at the
post-test phase. In all of the tests, it cannot be always proved that
negative evidence is more effective for parameter resetting, but it
seems likely that it is effective for learning transfer to take place.

The results that more teaching and transfer effect was observed
in the comparison test than the other two tests might be accounted
for by the feature of the test that the subjects had only to focus on
the syntactic difference between the two sentences presented.

NOTES

1) In L1 acquisition correction by a mother (negative evidence) does not influence the process of the child's restructuring grammar, as illustrated below.

Child: I putted the plates on the table.
Mother: You mean, I put the plates on the table.
Child: No, I putted them on all by myself.

(cited from Lightbown and Spada, 1993: 14)

2) The relationship between the subset hypothesis and the transfer hypothesis proposed by White (1989) has some similarities to the relationship between the creative construction hypothesis, which claims that "the second language growth is independent of a particular first language and develops rather in the manner in which a child 'creates' his first language" and the restructuring hypothesis, which claims that "a second language learner develops his second language by a process of restructuring his first language" (Stern, 1983: 396).

According to Dulay, Burt, and Krashen's analysis, only 4.7% of the children's errors could be classified as interlingual errors, on the other hand, 87.1% could be classified as intralingual errors. A number of similar results advocated the creative construction hypothesis (Dulay, Burt and Krashen, 1982).

3) Schwartz and Gubula-Ryzak (1992) pointed out that "the L2 behavior of these acquirers demonstrates that they have only learned that certain surface patterns are either possible or impossible in English" (p.4). White (1992) agrees with the comments and concludes that "the negative evidence supplied in White's original studies probably did not engage UG at all and that these learners learnt something superficial which they applied and later forgot" (p.136).

4) White (1989a; 1989b) used the strict adjacency parameter, while White
The Role of Positive and Negative Evidence in the Classroom

(1991) and Trahey and White (1993) replaced it with the verb-raising parameter in terms of the differences of adverb placement between English and French language. In order to compare adverb placement by Japanese and English language, I used the strict adjacency as a parameter.

5) Hatch and Farhady (1982) recommend that the time lag between the pretest and the post-test should be at least two weeks (p.22).

REFERENCES

Carter, R. "Language Awareness and English Language Teaching." Lecture at Temple University Japan. Osaka, 1994


Ellis, R. *Second Language Acquisition and Language Pedagogy*. Clevedon: Multilingual Matters, 1992


White, L. "The Adjacency Condition on Case Assignment: Do L2 Learners

——. *Universal Grammar and Second Language Acquisition.* Amsterdam: John Benjamins, 1989b.


### Appendix A

A: Information gap task

Sheet A

**DIRECTIONS:** Exchange the information you have and fill in the blanks with ○ or ×.

- 1. Peter opened the door
- 10. Fred often eats ice cream.
- 11. John drank his hot milk quickly.
- 16. Emi opened the window slowly

B: Consciousness-raising task

**Rule finding**

**DIRECTIONS:** Compare the Japanese sentences with the English ones, and find three rules on adverb placement.

2a. ジェーンは丁寧に部屋を掃除した。
2b. ジェーンは部屋を丁寧に掃除した。
2c. Jane *carefully* cleaned her room.
2d. *Jane cleaned *carefully* her room.

5a. ディックは時々友達に手紙を書く。
5b. ディックは友達に手紙を時々書く。
5c. Dick *sometimes* writes a letter to his friends.
5d. *Dick writes *sometimes* a letter to his friends.
Error correction

DIRECTIONS: Correct the following sentences.
6. Tom reads usually the newspaper.
8. John crossed carefully the road.
10. Fred eats often ice cream.
16. Emi opened slowly the window.

Appendix B

A: Grammaticality judgment test
DIRECTIONS: Choose one of the options; correct (a), incorrect (b).
1. John reads often the magazine,
2. Alice cut in the freezer some ice cream.
3. Susan usually plays tennis after school.
7. Mary played with his friend tennis.
10. Taro studied English with Jane.
14. Bill closed the door slowly.
15. Emi bought some flowers today.
16. Peter put the camera in the box.
19. Emi opened slowly the door.
21. My mother gave to me some advice.

B: Comparison test
DIRECTIONS: Choose one of the options; only a. is right (a); only b. is right (b); both right (c).
6. a. John drank his coffee quickly.
    b. John drank quickly his coffee.
7. a. John slowly ate ice cream.
    b. John ate ice cream slowly.
8. a. Harry quickly runs to his house.
    b. Harry runs quickly to his house.
9. a. Jane often eats ice cream.
b. Jane eats often ice cream.

10. a. Dick will buy tomorrow a new shirt.
   b. Tomorrow, Dick will buy a new shirt.

12. a. Fred is reading the newspaper in the kitchen.
   b. Fred is reading in the kitchen the newspaper.

15. a. Jane sometimes goes to the movies.
   b. Jane goes sometimes to the movies.

C: Manipulation test
DIRECTIONS: Rearrange the following words into correct order.

1. carefully, Jane, the problem, examined
2. always, Peter, black coffee, drinks
3. to me, Ellen, some chocolates, gave
4. slowly, John, to the park, walked
ABSTRACT

If the Subset Principle operates in L2 learning as it does in L1 acquisition, then L2 learners generate a subset grammar with positive evidence alone. In contrast to this view is a position which assumes that L2 learners no longer apply the Subset Principle, instead they will be influenced by their L1.

The purpose of this study is to investigate whether positive evidence alone can trigger parameter resetting, or negative evidence is indispensable to reset parameter value once set in L1 acquisition and to seek what kind of teaching can help Japanese learners parameter-reset effectively.

Japanese and English are in a superset/subset relation in the case of adverb placement, since Japanese adverbs can intervene between a verb and its direct object, while English counterpart cannot, which is referred to as Adjacency Condition on Case assignment. Adapting parametric variation, English is [+strict adjacency] and Japanese is [−strict adjacency].

The study involved the two kinds of treatment as the experimental conditions. Information gap activities were given to a group, which intensively provided positive evidence as input. Another class included
rule-finding and error correction activities, which provided negative
evidence as input. Teaching focused on the positions of the two
kinds of English adverb: adverbs of manner and adverbs of frequency.

Grammaticality judgement test, comparison test, and manipulation
test were administered to measure both direct and indirect effect of
learning.

The results show that not only the negative evidence group but
also positive evidence group performed better in the post-test than in
the pre-test, although the learning effect declined at the final stage.
A learning transfer was observed. The negative evidence group
distinguished SVPPO error sentences from grammatical sentences bett-
er than the positive evidence group.