Shintani, N. (2013). The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of L2 vocabulary by young beginning-level learners. *TESOL Quarterly, 47*(1), 36-62. doi: 10.1002/tesq.54

The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of L2 vocabulary by young beginner learners

Natsuko Shintani

University of Auckland

Introduction

The purpose of the study reported in this article is to investigate the effect of two instructional approaches–focus on forms (FonFs) and focus on form (FonF)–on the acquisition of a set of nouns and adjectives by young Japanese children who were complete beginners. I begin by defining FonFs and FonF and consider the theoretical rationale for each. I then examine how focus on forms (operationalized as "present-practice-produce") and focus on form (operationalized through "task-based teaching") can be realized. I will point out that in the case of beginner learners "task-based teaching" is best operationalized in terms of comprehension-based rather than production-based tasks. I conclude with a review of the research that has compared the effects of FonFs and FonF on vocabulary acquisition.

Definitions

In "Focus on forms (FonFs)" (Long, 1991, 1996) language is broken down into discrete elements (e.g., words, grammar rules, notions or functions), which are then taught item by item in a linear, additive fashion. FonFs, therefore, constitutes a traditional approach to language teaching involving a linear syllabus, instructional materials and corresponding procedures designed to present and practice a series of linguistic items. In this type of instruction, the learners' primary attention is directed at linguistic form but meaning is not excluded.

In "Focus on form" (Long, 1991; Long & Crookes, 1992) the primary focus is on meaning (i.e., on message processing) rather than on form. It involves an occasional shift of learner's attention from meaning to a linguistic form and the meaning this conveys while the overriding focus remains on communicating. This shift can be triggered by perceived problems with either comprehension or production and it can be initiated by either the teacher or students. A key feature of FonF instruction is that it emphasizes form-function mapping.

Ellis, Basturkmen and Loewen (2002) pointed out that FonF can be planned or unplanned. In planned FonF the focus on a specific linguistic feature is pre-determined and a "focused task" (p. 420) is designed to provide a context for its use. Unplanned FonF takes place when the learners' attention is directed incidentally to specific linguistic forms while they are performing unfocused tasks. Thus, in the case of unplanned FonF, the linguistic forms

that are attended to are not selected in advance but arise naturally in the performance of a communicative task. In planned FonF, attention to the selected linguistic feature is intensive (i.e., attention is directed continuously at the same feature) while in unplanned focus-on-form, attention to form is usually extensive (i.e., it is directed at a variety of different linguistic features).

As defined by Ellis, Basturkmen and Loewen (2002), FonF and FonFs instruction involve incidental and intentional learning respectively. Intentional acquisition takes place when the learners function as learners and consciously seek to learn some specific feature. Incidental acquisition occurs when learners function as communicators and learn the L2 without any intention to do so. However, as Schmidt (1994) pointed out, incidental acquisition may require conscious attention to linguistic form (i.e., noticing). Thus, what distinguishes incidental and intentional language learning is not consciousness as noticing but the lack of any intention to learn.

Theoretical Background

FonFs instruction draws on a number of theoretical positions regarding how an L2 is learned. Recently, it has drawn on Anderson's (1983) Skill Acquisition Theory for support. DeKeyser (2007), for example, proposes that L2 acquisition can be best achieved by transforming declarative knowledge into procedural knowledge through "practice". DeKeyser argues that the process of second language learning starts with declarative knowledge (i.e., understanding particular linguistic features) which is then transformed into procedural knowledge through intensive practice. He defines "practice" as "specific activities in the second language, engaged in systematically, deliberately, with the goal of developing knowledge of and skills in the second language" (p.1). DeKeyser emphasizes that drill-like practice is not enough; learners also need to experience using the L2 under real operating conditions (i.e., in the performance of a communicative task). However, this does not amount to "FonF" as teaching is based on a structural syllabus and involves intentional learning.

FonF instruction draws on current theories that emphasize the role of interaction and cognitive processing in L2 acquisition. Two key interactional constructs are the "negotiation of meaning" and the "negotiation of form". These serve to draw learners' attention to form while they are communicating. They induce the cognitive processes of "noticing" and "noticing the gap".

Negotiation of meaning takes place through the collaborative work that speakers undertake to achieve mutual understanding when there is some kind of communication problem (Long, 1983). Interactionally modified input promotes acquisition in two ways: (1) when learners notice the "positive evidence" (Long, 1996) provided in the input and when they notice the gap between this and their own interlanguage and (2) when negative evidence and modified output occur as a result of corrective feedback following learners' erroneous production.

A number of studies have investigated the effect of the negotiation of meaning on L2 vocabulary acquisition by comparing "interactionally modified input" (i.e., the learners were able to interact to negotiate the meaning of the input) with "baseline input" (i.e., the learners received unmodified input and were unable to negotiate) and "premodified input"

(i.e., the learners received simplified input and were unable to negotiate). For example, Ellis, Tanaka and Yamazaki (1994) investigated the effects of modified interaction on the comprehension and vocabulary acquisition of high-school students of English in Japan. The results indicated that interactionally modified input resulted in better comprehension and more new words being acquired than either baseline or premodified input. Mackey and Goo's (2007) meta-analysis shows that when learners have the opportunity to negotiate for meaning, acquisition can take place

FonF can also take place even when there is no communication breakdown. As Ellis, Basturkmen, and Loewen (2001) pointed out, in classrooms it is common for negotiation to occur even though no communication problem has arisen. In this case, the focus is on a linguistic problem not a communicative problem. A typical example is when the teacher corrects a learner's erroneous production even when there is no communication problem. For example, if a learner said "I go cinema yesterday," the teacher would know that the learner is referring to the past but might still correct using a recast ("oh you went to cinema yesterday"). Ellis et al. (2001) refer to this as the "negotiation of form". They point out that such negotiation can still be considered a type of FonF because, like the negotiation of meaning, it arises when learners are attempting to communicate.

Pedagogic Realizations of FonFs and FonF

In many current instructional materials, FonFs is realized in terms of present-practice-produce (PPP; Ur, 1996). DeKeyser argues that such an approach is ideally suited to older learners who have lost the ability for the kind of implicit learning that children are capable of. However, PPP also figures strongly in instructional materials for children including those who are complete beginners (e.g., Nakata, Frazier, Hoskins, & Graham, 2007). A key feature of PPP is that it seeks to elicit production of correct target forms right from the start as a means for learning them. PPP includes meaning-based activities as well as controlled production exercises but when learners perform them, they are likely to be aware that the purpose is not to "communicate" but to "practice" specific linguistic forms.

FonF can involve a variety of instructional activities. Doughty and Williams (1998) distinguish them in terms of the extent to which they are unobtrusive or obtrusive "reflecting the degree to which the focus on form interrupts the flow of communication" (p. 258). Thus, "input flood" and "task-essential language" constitute relatively unobtrusive types of FonF whereas "consciousness-raising" and "input processing" are obtrusive. These different types of FonF also differ in terms of whether they involve reactive or proactive attention to form. For example, FonF involving "tasks" will entail the use of reactive techniques that induce "on-the-spot" attention to form as the task is performed. In contrast, "consciousness-raising" activities are proactive as they focus on features that learners are made explicitly aware of. In the study reported in this article, FonF has been operationalized as unobtrusive and reactive.

FonF of this kind requires a task-based approach. Ellis (2003) proposed four defining criteria for a "task": (1) meaning is primary; (2) there is some type of gap (e.g., an information gap); (3) learners are required to use their own linguistic and non-linguistic resources to communicate, and (4) there is some outcome other than simply the display of correct language. Tasks can be input-based or output-based. In input-based tasks, learners

are asked to comprehend input in order to achieve the task outcome, which involves some kind of non-verbal response (e.g., choosing the correct picture). They can be designed in such a way that learners are only able to achieve the outcome if they have both "noticed" and comprehended the specific linguistic forms needed to achieve the outcome. Learners receive feedback on their non-verbal responses and this plays an important role in enabling them to see if they have processed the input correctly. Opportunity for focus of form also arises when there is negotiation of meaning or form. Ellis (2003) suggests that "simple listening tasks can be devised that can be performed with zero competence in the L2" (p. 37).

Comparative Vocabulary Studies of FonF and FonFs

Despite the interest in the relative effectiveness of FonFs and FonF, there have been only a few studies that have explicitly compared the two approaches for vocabulary acquisition.

Laufer (2006), distinguished the two types of instruction as follows: FonF "views the words attended to as tools for task completion, and the FonFs approach treats the words attended to as the objects of study" (p.150). She conducted a study comparing the effect of comprehension-based FonF (i.e., reading tasks) and production-based FonFs (production exercises) on the acquisition of new English vocabulary by Grade 11 students in Israel. The learners in the FonF condition were asked to read a text and answer comprehension questions using a bilingual dictionary if they needed. This condition catered to incidental language acquisition. In the FonFs condition the learners first received a list of target words with translations and explanations in English, and the teacher also explained each word orally. Then the learners worked on word production exercises. This catered to intentional language learning. Acquisition was measured by a discrete-item comprehension test (providing L1 translations or L2 definitions of the target words). The results showed that FonFs group outperformed the FonF group. Laufer argued that learners do not necessarily notice unfamiliar words in input and thus "the nature of lexical competence makes FonFs indispensable to vocabulary instruction." She also suggested that FonFs is of "major importance in any learning context that cannot recreate the input conditions of first-language acquisition" (p. 162). Laufer and Rozovski-Roitblat (2001) obtained very similar results in a follow-up study.

The findings of Laufer's studies suggest that FonFs (intentional learning) is more effective than FonF (incidental learning) for vocabulary acquisition. This might be because FonFs forces learners to "notice" the forms of words. However, there are number of studies that show that FonF instruction can also provide opportunities for learners to "notice" word form (e.g., Gass & Torres, 2005). The key issue, then, might be to what extent the FonF leads to "noticing" form-meaning connections. It could also be argued that incidental acquisition is more likely to occur in young beginner learners of L2 than intentional learning. Thus, two questions arise here: (1) whether FonFs is always superior to FonF in inducing "noticing" of vocabulary items and (2) whether intentional acquisition is superior to incidental acquisition for young beginner learners.

De la Fuente (2006) provided an answer to the first question by comparing PPP (FonFs) with production-based TBLT (FonF) with university students in an elementary Spanish class. The PPP learners received 50 minutes of instruction consisting of explanation of the

new words (presentation), controlled oral and written production exercises (practice) and a role play performed in pairs (free production). The students in the two TBLT conditions worked on a restaurant task where students need to negotiate the meaning of the target words to complete the task. The acquisition was measured by means of a discrete-item oral production test. The results showed that the two conditions were equally effective in the immediate post-test. However, in the delayed post-test, the TBLT group outperformed the PPP group. Examining the interactions that occurred in each group, de la Fuente found that the TBLT instruction provided more opportunities for the negotiation of meaning, production of the target words, and on-line retrieval of target words than the PPP. However, a limitation of the study was that acquisition was only measured by means of a discrete-point test.

Shintani's (2011) study attempted to answer both of the above questions. She compared the effects of input-based tasks (FonF) with production-based activities (FonFs) on the acquisition of a set of English concrete nouns by young beginner learners. The input-based groups received a set of listen-and-do tasks which required the learners to identify and comprehend the target nouns. The production-based group received oral production-based activities in accordance with the three phases of PPP. Acquisition was measured by four tests (comprehension/ production in both discrete-item and task-based conditions). The results showed that both input-based tasks and production-based activities were effective in developing both receptive and productive vocabulary knowledge. Interestingly, Shintani found similar levels of productive knowledge in the two groups despite the significantly fewer opportunities for output in the input-based group. She explained the results in terms of differences in the nature of interactions that occurred in the two groups. While the interactions in the production-based group consisted or IRF (i.e., initiate-respond-feedback) exchanges (Sinclair & Coulthard, 1975), the interactions in the input-based group involved negotiation of meaning initiated by the students.

These studies have produced mixed results. This is not surprising as they operationalized FonF and FonFs instruction in different ways. A key difference between Laufer's studies and the other two studies lies in whether the learners had the opportunity to engage in interaction where they could "negotiate" the target items in the FonF instruction. De la Fuente's and Shintani's studies suggest that when such an opportunity is available, FonF proves as or more effective than FonFs.

The study reported below aimed to further investigate the comparative effectiveness of input-based FonF and production-based FonFs on vocabulary acquisition. It extended the study by Shintani (2011) by including adjectives as well as nouns as the target items. It also examined both the process features of the two kinds of instruction and the learning products. [1]

Method

Research Questions

RQ 1.To what extent is the difference between FonFs and FonF reflected in the process features of the two types of instruction?

RQ 2.What effect does FonFs have on young, beginner learners' productive knowledge of nouns and adjectives? [2]

RQ 3. What effect does FonF have on young beginner learners' productive knowledge of nouns and adjectives?

RQ 4. Is there any difference in the effect of FonFs and FonF on young beginner learners' productive knowledge of nouns and adjectives?

Research Design

Materials for three different lessons were designed for the FonFs, the FonF and the control groups. Each lesson lasted approximately 30 minutes. The lesson for each group was repeated nine times over five weeks but the FonFs group was exposed to different sets of words each time as explained below. The acquisition of productive knowledge was measured by a discrete-point test and a task-based test, which were conducted as pre-test (one week before the treatment), post-test 1 (one week after the treatment) and post-test 2 (four weeks after the treatment). All the lessons were taught by the researcher, who had ten years of teaching experience.

Target Items

The previous study (Shintani, 2011) showed that the young beginner learners successfully acquired nouns through both FonFs and FonF instruction. This study included adjectives as well as nouns as the target words because, as Gasser and Smith (1998) noted, adjectives are more difficult to learn than nouns for children.

24 nouns and 12 adjectives were chosen as the target words. Three criteria guided the selection of the target words: (1) the L1 equivalents of the words would be known by the Japanese children, (2) the words were not in the *List of English Loanwords in Japanese* (Daulton, 1999), avoiding the possibility of participants inferring the word meanings by using their L1 knowledge and (3), in the case of the adjectives, they served as common descriptors of the target nouns. The 24 nouns met criteria (1) and (2). However, most of the adjectives did appear in Daulton's list (the exceptions being *purple*, *small*, and *short*) because of the importance of criterion (3). A list of the target items is provided in Appendix A in the online Supporting Information.

One characteristic of FonFs is that it employs a linear syllabus. In order to operationalize this, the target items were divided into three sets (Set A, B and C) and introduced in different lessons for the FonFs group. Each set consisted of eight nouns and four adjectives. Set A was introduced in lessons 1 and 2, set B in lessons 3 and 4, and set C in lessons 5 and 6. Then each set was taught once again in lessons 7, 8 and 9 respectively as review lessons. In order to balance the time difference between exposure to the target items and the post-tests, the FonFs group practised the previous target items in each lesson. In the FonFs both nouns and adjectives were introduced in accordance with PPP methodology; that is, the form and meaning of the words were introduced explicitly before production practice.

In the FonF group, on the other hand, all the target words appeared incidentally in every lesson as FonF instruction typically involves exposure to a range of different language features (Ellis et al., 2002). The nouns and adjectives appeared in different ways. The learners had to identify picture cards that corresponded to the nouns in the teacher's

direction. Each direction involved one noun and all the 24 nouns appeared in every lesson. On the other hand, the adjectives arose only spontaneously in the process of performing the tasks. Therefore, the 12 adjectives appeared only randomly in the nine lessons. As will be reported later, however, the teacher's input contained many exemplars of the target adjectives.

Participants

This study involved 45 Japanese children aged six with no experience of L2 learning. That is, they were complete beginners of English. The participants were randomly divided into three groups of 15: two experimental groups (i.e., the FonF and FonFs groups) and one control group. Each group was further divided into two classes of between six to nine participants. All groups met twice a week during the project and did not have any other English instruction for the duration of this study.

Instructional Treatments

Instructional materials and procedures for the FonFs group. Prior to each FonFs lesson, the goal of the activities–to learn new vocabulary–was made explicit to the students to induce intentional learning of the words. Each lesson then consisted of five activities representing the three phases of the FonFs (present-practice-produce). The first activity served as the "present" phase, involving the participants' repeating individual words. The second and the third activities, which served as the "practice" phase, involved the participants' saying the word shown on flash cards–both chorally and individually. The fourth and the fifth activities served as the "produce" (i.e., free production) phase. For example, one such activity involved the students choosing a card from a set of 30 cards face down on the table in turn. If a student could say the word for the object shown on the card and if a card showing the same object was face up on the table, the student could collect the two cards. If not, the card the student had chosen was left face-up on the table. Each student counted the cards he/she had collected and the student with the most cards was the winner. Opportunities for learning vocabulary occurred when the learners produced the target words and received corrective feedback from the teacher. Throughout the five activities, the students were required to demonstrate accurate production [3]. All five activities were carried out in each lesson and repeated nine times with different sets of items. The researcher used English during the activities but the learners' first language (Japanese) was used to explain the procedures for each activity whenever necessary.

Instructional materials and procedures for the FonF group. Three tasks were designed in accordance with Ellis's (2003) definition of tasks (see earlier section). The tasks were designed in such a way that the outcome could only be achieved if the learners were successful in comprehending the input. Each task involved the participants' listening to the teacher's commands (e.g., "Please take the crocodile to the zoo") and responding to them (e.g., by choosing the correct card and placing it in the correct holder).

At the beginning of each lesson, the goal and task procedures were explained to the participants using Japanese. However, the instruction was mostly conducted in English.

Instructional materials for the control group. The control group attended regular lessons in the private school. The lessons consisted of practicing English songs, Total Physical Response (Asher, 1977), and tracing and copying the alphabet on worksheets. The

control group was taught by the same teacher as the experimental groups. Care was taken to avoid using any of the target words with the control group.

Testing Materials

In order to reduce test bias, two production tests (i.e., a task-based and a discrete-item test) were designed [4].

Discrete-item word production test. The test required individual participants to name the target vocabulary items on each flash-card. Twenty-four flash cards representing the target nouns were used. The researcher elicited the learners' production by saying "what's this?" in English. If a learner did not understand the question, the researcher used Japanese to help out. For the adjectives the six colours appeared on flash cards. The researcher asked "what colour is this?" in English or Japanese to elicit production. The cards for the dimension adjectives were different from the ones used in the FonFs lessons in order to avoid bias in favour of this type of instruction. The researcher only used Japanese (e.g., *"this box is __?"* in Japanese) to elicit production (i.e., she did not use the words introduced in the instruction). The adjectives were tested after the noun items.

Same or Different Task test. The students performed the task one-on-one with the researcher. Each had a sheet with 24 numbered pictures depicting the target nouns and adjectives. The researcher's and the students' sheets contained matching and non-matching pictures. The purpose of the task was for the students to determine whether their pictures were the same as or different from the teacher's. The researcher asked questions to elicit the students' production by asking questions such as "What colour is it?" or "My soap is pink. Is your soap pink?" The interactions were audio-recorded. The recordings were then checked to see if a student had correctly produced the target words.

For both tests, one point was awarded for each item that a student produced correctly. The total possible was 24 for nouns and 12 for adjectives. The researcher scored the test while she was conducting it [5].

As all the participants had only just entered primary school, it was unlikely that any of them were familiar with the procedure for either test.

Data Analysis

The nine lessons for the FonFs and FonF groups were audio- and video-recorded and transcribed. The video-data were used to identify the individual learner utterances and non-verbal responses. The transcribed data were first analysed in terms of the number of tokens of the target items in the input and output that occurred in the two experimental groups. A student token was counted as only one if more than one student chorally stated a word. As the students were exposed to the target language produced by other students as well as the teacher, student tokens were added to the teacher tokens to estimate the total amount of input. The transcripts were further analysed in terms of the occurrence of different types of input and output of the target words (see below).

The test scores were first analysed in terms of descriptive statistics, within- and between-group differences using PASW version 18. A series of non-parametric tests were applied as the scores for the nouns did not meet the assumption of normality. For the adjective tests, a series of parametric analyses were undertaken. Categorical data were analysed using the

chi-square test. Effect sizes for the comparative effects of two variables were estimated using the correlation coefficient effect size *r*. For the chi-square tests, Cohen's *w* was used. Effect size values were interpreted as small = .1, medium = .3 or large = .5 as Cohen (1988) suggested.

Results

The process features of the two types of instruction will be examined first. Then the results for nouns will be presented followed by those for adjectives.

Process Features

Frequency of target words in input and output. The total amount of input and output in the FonFs and FonF groups is shown in Table 1.

Table 1

Frequency of Target Words in Input and Output

	Input			Output		
	FonFs	FonF	Difference	FonFs	FonF	Difference
Nouns	6,030	5,940	n.s.*	2,766	52	sig.*
Adjectives	2,442	2,520	n.s*	1,080	301	sig.*

Note: * n.s. = non significant, sig. = significant at .01 level.

Both groups were exposed to substantial input. However, the opportunity for output was somewhat different in the two groups. Chi-square tests showed that there were no significant differences between the two groups in terms of input for both nouns and adjectives (nouns: $\chi^2 = .677$, $df = 1$, $p > .01$, $w = .01$; adjectives: $\chi^2 = 1.226$, $df = 1$, $p > .01$, $w = .01$). However, there were significant differences in output for both nouns and adjectives (nouns: $\chi^2 = 2613.838$, $df = 1$, $p > .01$, $w = .96$; adjectives: $\chi^2 = 439.421$, $df = 1$, $p > .01$, $w = .56$). In other words, both groups were exposed to a similar number of input tokens but the FonFs group had more opportunity to produce both the nouns and adjectives than the FonF group.

Characteristics of input. In order to compare the differences in input, the teacher's target word tokens were categorized as "embedded" or "isolated". When the teacher used a word in a longer utterance, it was coded as "embedded". If the teacher produced a word by itself, it was coded as "isolated."

In the FonFs lessons, the teacher often produced isolated words when providing feedback on the learners' erroneous production as in Excerpt 1.

Excerpt 1

1. T: (pointing to the shorter pencil in the picture card), this pencil is…?

2. S: small.

3. T: short.

4. S: short.

(FonFs: Lesson 3)

In the FonF group, on the other hand, the target words were frequently embedded in longer utterances. Excerpt 2 illustrates how they often appeared in the teacher's commands and also in the descriptions of the target nouns when the teacher tried to help the learners select the appropriate noun cards. In both cases, the target items were embedded in complete utterances.

Excerpt 2

1. T: Please take the peacock, peacock to the zoo.

2. S: Peacock.

3. T: Yeah. The peacock has a beautiful feather. It's a blue bird. Blue.

4. S: Blue.

5: T: Yeah. The peacock is blue. Okay? Ready?

(FonF; Lesson 3)

As Table 2 shows, the FonF group was exposed to embedded words more frequently than the FonFs group.

Table 2

"Embedded" and "Isolated" Word Production by the Teacher

	FonFs		FonF	
	Nouns	Adjectives	Nouns	Adjectives
Embedded	152	52	3,991	2,057
Isolated	3,112	1,310	1,897	162

Chi-Square tests showed that for both nouns and adjectives there were significant differences between the two groups in the number of the "embedded" word tokens (nouns: $\chi^2 = 3557.307$, $df = 1$, $p<.01$, $w = .93$, adjectives: $\chi^2 = 1906.128$, $df = 1$, $p<.01$, $w = .95$) and in the "isolated" tokens (nouns: $\chi^2 = 294.715$, $df = 1$, $p<.01$, $w = .24$, adjectives: $\chi^2 = 895.315$, $df = 1$, $p<.01$, $w = .78$). In other words, the FonF learners experienced the target words–both noun and adjectives–in an embedded form whereas the FonFs learners experienced them as isolated words.

Cross-tabulation analyses showed a significant difference between the nouns and adjectives in terms of the proportion of "embedded" and "isolated" tokens in the FonF

group (χ^2 = 528.111, *df* = 1, *p*<.01, *w* = .26) but no significant differences in the FonFs group (χ^2 = 1.605, *df* = 1, *p*>.01, *w* = .02). In other words, in the FonF, the teacher was more likely to use adjectives in an embedded manner than nouns whereas in the FonFs group both nouns and adjectives were produced in a similar way.

Characteristics of learner production. The characteristics of learner production were examined by categorising the students' target word tokens as "optional" or "requested". Production was deemed to be "optional" if the teacher had not made an attempt to elicit a target word from the students. Production was considered "requested" whenever there was an explicit attempt to elicit a target word. For example, in Excerpt 1 (from the FonFs group), the student's production of "small" (line 2) was requested by the teacher but "*short*" (line 4) was optional. Further examples of optional production of the target forms can be seen in lines 2 and 4 in Excerpt 2.

All the above examples of "optional" production consisted of repetitions of a part of the teacher's utterances. However, there were a number of occasions when learners produced words independently. This was particularly common for adjectives in the FonF group as shown in Excerpt 3. The turn in line 8 is the repetition of the teacher's utterance but in the turns in lines 2, 4, and 6 the students initiate the production of adjectives.

Excerpt 3

1. T: please take the ostrich to the zoo.

2. S1: <u>green?</u>

3. T: not green.

4. S1: <u>blue?</u>

5. T: no.

6. S1: <u>gold?</u>

7. T: gold? no. the ostrich is big.

(FonF: Lesson 5)

In order to examine this, the "optional" tokens were further classified as "self-initiated" and "borrowed". A target word was considered to be initiated if the learner was the first person in a sequence to use it. It was considered to be borrowed if the target word had occurred in an utterance earlier in a sequence and the learner was just repeating or imitating it. Table 3 shows the number of the adjective and noun tokens produced by the two groups in each of the categories "requested/optional" and "self-initiated/ borrowed" in the nine lessons.

Table 3

Number of Requested/ Optional and Self-Initiated/ Borrowed Tokens by the Students

	FonFs		FonF	
	Nouns	Adjectives	Nouns	Adjectives

Requested		2,601	1,028	0	0
Optional	Self-initiated	6	3	2	263
	Borrowed	159	49	50	38

The table shows that no "requested" production occurred in the FonF group. The FonFs group produced substantial output but it was because the production was "requested". The "optional" production in the FonFs group was mostly "borrowed" whereas the FonF students often "self-initiated", mainly for the adjectives. Cross-tabulation analyses showed a significant difference between the nouns and adjectives in terms of the number of "self-initiated" and "borrowed" tokens in the FonF group ($\chi^2 = 165.308$, $df = 1$, $p<.01$, $w = .68$) but this difference was non-significant in the FonFs group ($\chi^2 = .452$, $df = 1$, $p>.01$, $w = .05$).

Acquisition of Productive Knowledge

Acquisition of productive knowledge was measured by two production tests (i.e., the Same or Different task and the discrete-item production test). The detailed statistical results are provided in Appendix B in the online Supporting Information.

Nouns. Table 4 shows the descriptive statistics of the test scores for the nouns on the four tests. The maximum score possible for each test was 24.

Table 4

Descriptive Statistics for the Target Nouns

Test	Group	Pre-test		Post-test 1		Post-test 2	
		Mean	SD	Mean	SD	Mean	SD
Same or Different Task	FonFs (n=15)	0.20	0.56	7.87	4.05	7.20	3.97
	FonF (n=15)	0.20	0.56	5.07	2.52	6.33	2.74
	Control (n=15)	0.13	0.35	0.20	0.56	0.20	0.56
Discrete-item test	FonFs (n=15)	0.20	0.56	8.40	3.74	8.93	4.20
	FonF (n=15)	0.20	0.56	6.53	2.64	7.93	2.02
	Control (n=15)	0.20	0.56	0.27	0.59	0.40	0.74

Within-group comparisons showed that both the FonFs and the FonF groups significantly improved in post-test 1 and post-test 2 compared to the pre-test in both tests. However, while the FonFs group did not improve from post-test 1 to post-test 2 in either test, the FonF group significantly improved from post-test 1 to post-test 2 in both tests. The control group did not show any significant changes between the three repeated tests.

Between-group comparisons showed there was no significant difference between the three groups in pre-test for both the Same or Different Task and the discrete-item production test. In the post-tests, however, both experimental groups outperformed the control group in

both the Same or Different Task and the discrete-item test. However, no significant differences were found between the two experimental groups either in post-test 1 or in post-test 2 in either test.

Table 5 summarises the comparative effect sizes of the three groups for the acquisition of nouns.

Table 5

Comparative Effect Sizes (*r*) for the Acquisition of Nouns

Comparative results	FonFs>Cntl.		FonF>Cntl.		FonFs>FonF	
	Post-test 1	Post-test 2	Post-test 1	Post-test 2	Post-test 1	Post-test 2
Same or Different	.80*	.78*	.80*	.84*	.38	.13
Discrete-item	.84*	.82*	.85*	.82*	.28	.15

Note: Cntl.; Control group

Note: *the difference was significant at p=0.05 level.

Both experimental groups showed large effect sizes when they were compared with the control group. Although the FonFs group scored higher than the FonF, the effect sizes were either negligible or small, and none of the two groups' post-test scores were significantly different. In other words, the FonF and the FonFs instruction had a very similar effect.

Adjectives. Table 6 shows the descriptive statistics for the target adjectives. The maximum score possible for all tests was 12.

Table 6

Descriptive Statistics for the Target Adjectives

Test	Group	Pre-test		Post-test 1		Post-test 2	
		Mean	SD	Mean	SD	Mean	SD
Same or Different Task	FonFs (n=15)	0.67	1.45	1.67	1.80	1.67	1.76
	FonF (n=15)	0.60	1.30	4.47	2.77	4.87	3.14
	Control (n=15)	0.60	1.24	0.73	1.58	0.67	1.40
Discrete-item production test	FonF (n=15)	1.73	1.53	5.53	2.95	6.53	2.75
	FonF (n=15)	1.73	1.53	5.53	2.95	6.53	2.75
	Control (n=15)	1.40	0.91	1.13	0.99	1.07	0.96

Within-group comparisons showed a significant test effect for the Same or Different Task. Pairwise comparisons showed that the FonFs significantly improved from pre-test to post-test 1 but did not improve from pre-test to post-test 2 or from post-test 1 to post-test 2.The

FonF group, on the other hand, showed significant improvement from pre-test to both post-tests but no difference was found between post-test 1 and 2. A significant test effect was also found in the discrete-item test. In this test, both experimental groups improved significantly from pre-test to both post-tests, but there was no significant difference between the two post-tests. The control group did not show any significant changes over time in either test.

Between-group comparisons showed there was no significant effect in the pre-test but there were significant effects in both post-tests 1 and 2 for both the Same or Different Task and the discrete-item test. The pairwise comparisons showed that the FonF group significantly outperformed both the FonFs and the control groups in both tests. Although the FonFs group significantly outperformed the control group in the discrete-item test in the post-tests, this group failed to outperform the control group in the Same or Different tasks in either post-test.

Table 7 summarises the comparative effect sizes of the three groups for the acquisition of adjectives.

Table 7

Comparative Effect Sizes (*r*) for the Acquisition of Adjectives

Comparative results	FonFs>Cntl.		FonF>Cntl.		FonF>FonFs	
	Post-test 1	Post-test 2	Post-test 1	Post-test 2	Post-test 1	Post-test 2
Same or Different	.33*	.30	.63*	.65*	.51*	.53*
Discrete-item	.50*	.50*	.71*	.80*	.39*	.58*

Note: Cntl.; Control group

Note: difference was significant at $p = 0.05$ level.

The table shows that the FonFs group's test scores improved but the effect sizes were smaller than those for the FonF group. In particular, the FonFs group failed to show a significant gain in comparison to the control group in post-test 2 for the Same or Different task. The FonF significantly outperformed the FonFs with small to medium effect sizes.

Discussion

There were two reasons for investigating the process features of the FonFs and FonF instruction. One was to demonstrate that the "internal" characteristics of the two types of instruction (i.e., the activity that arises in the classroom) matched their "external" descriptions (i.e., the activity predicted by the design of the instruction). The second was to identify process features that might help to explain differences in the experimental groups' learning outcomes.

The FonFs instruction taught the target forms directly. It catered to intentional learning and emphasized production. The FonF instruction aimed to draw attention to form while the learners were primarily focused on meaning. It catered to incidental learning by asking the

learners to perform tasks. Given that the learners in this study were complete beginners the tasks were all input-based although learner production was not prohibited. Table 8 shows the key differences in the process features of the two types of instruction.

Table 8

Differences in the Process Features of the FonFs and FonF Groups

Process feature	FonFs instruction	FonF instruction
Input		
Frequency of exposure to target items	Frequent exposure to the target items	Frequent exposure to target items
Isolated versus embedded	Target items generally occurred in isolation.	Target items often embedded in utterances. The adjectives were more often embedded than the nouns.
Output		
Frequency of production of target items	Frequent opportunities for production of the target words	Fewer opportunities to produce the target words
Optional vs. requested	The production of the target words was mostly requested in IRF exchanges.	The production of the target words only occurred optionally.
Initiated vs. borrowed	Learners produced the nouns and adjectives by borrowing them from the teacher's utterances.	The production of nouns generally occurred by borrowing from the teacher's utterances. However, the production of adjectives was often initiated by the students in order to negotiate the meaning of the nouns.

Table 8 demonstrates that there were marked differences in the two kinds of instruction and also that the differences reflect the fundamental nature of FonFs and FonF. Thus, as might be expected in lessons based on PPP, the learners were frequently exposed to the target words but almost always in an isolated manner. They typically produced the target words when requested to do so by the teacher and they only rarely initiated production themselves. The teacher's request usually triggered the three part exchange (i.e., initiation-response-follow-up, or IRF; Sinclair & Coulthard, 1975) so typical of FonFs instruction–as shown in Excerpt 1. As other studies (Brock, 1986; Long & Sato, 1983) have shown, these IRF exchanges involve display questions (i.e., questions where the answers are already known to the person asking them) rather than referential questions (i.e., questions where the answers are not known by the person asking them), as again shown in Excerpt 1. Furthermore, when the FonFs learners produced the target words, they frequently "borrowed" them from the teacher's corrective feedback move following their own erroneous production. These process features suggest that the FonFs learners were treating the target words as objects that had to be remembered.

The FonF group was also frequently exposed to the target words as might be expected in an approach based on input-based tasks. However, the nature of this exposure was very different from that of the FonFs group. The target words were frequently embedded in the teacher's complete sentences. Thus, the FonF instruction exposed the learners to the target words in a contextualized and naturalistic way. When the learners produced the target words, they did so because they had opted to produce them rather than because they were required to do so, reflecting the nature of the comprehension-based instruction. Interestingly, the adjectives were embedded more frequently than the nouns. This was because the learners engaged in the negotiation of meaning by initiating use of the adjectives in order to comprehend the teacher's commands (see Excerpt 3). In other words, the FonF instruction created a situation where the learners needed to use their own linguistic resources. All of these features are characteristics of the kind of language use associated with focus-on-form.

It is clear that the FonFs and FonF instruction resulted in markedly different processes in ways to be expected from the external descriptions of these two types of instruction.

Research question 2 asked whether FonFs instruction was effective in developing the learners' productive knowledge of both nouns and adjectives. The results showed that the FonFs group demonstrated significant productive knowledge of the nouns in both the discrete item and the Same or Different Task and in both the immediate and delayed version of these tests. In other words the instruction was effective in enabling the learners not only to use the nouns in a controlled production test but also a free production test. This result lends support to the claims of skill-learning theory, namely that explicit presentation of linguistic items followed by controlled and free production practice leads to the ability to use the items in communication.

However, the results for adjectives suggest that such instruction is not always effective. The FonFs group was only able to use the adjectives in the controlled production test. It failed to demonstrate that it had developed any communicative control over them. Also, as

shown in Table 7, the tests for the adjectives showed smaller effect sizes than for the nouns. Given that the process characteristics of the FonFs instruction were the same for both nouns and adjectives (see Table 8), it would seem that the explanation for the difference in the effects of this kind of instruction on the two word types was due to the greater learning difficulty posed by adjectives (Gasser & Smith, 1998). The processes involved in the FonFs instruction did not require much depth of processing. They were sufficient to enable the learners to acquire words for labelling concrete objects but not for acquiring the words needed to describe the qualities of these objects.

Research question 3 asked whether FonF instruction was effective in developing the learners' productive knowledge of both nouns and adjectives. The results showed that the FonF group developed productive knowledge of both nouns and adjectives in both the discrete point test and the Same or Different task (i.e. in free as well as controlled production) even though the instruction did not require them to produce the target words. It appears that learners do not need to practice producing words in order to develop productive knowledge of them. Shintani (2011) argued that comprehension-based instruction enables learners to develop productive knowledge of nouns because it provides opportunities for learners to negotiate their meaning, a key feature of FonF instruction. As shown in Table 3, many of the students engaged in the negotiation of the meaning of the nouns introduced in the teacher's commands (see Excerpt 3 for example). Thus, plentiful exposure to new vocabulary in a communicative task plus the opportunity to negotiate the meanings of the words when necessary seems sufficient for the development of productive knowledge.

In the FonF instruction the nouns were introduced systematically in the teacher's commands. However, the adjectives were only introduced when the learners experienced difficulty in identifying the referents of the nouns (i.e., they were taught reactively). In this respect the FonF instruction differed from the FonFs instruction. It would seem then that task-based teaching can foster acquisition both pre-emptively (in the case of the nouns) and reactively (in the case of the adjectives) as claimed, for example, Willis & Willis (2007).

Research question 4 asked whether there was any difference in the effects of FonFs and FonF instruction on the acquisition of productive knowledge of nouns and adjectives. The key difference in the results for the FonFs and FonF groups was that whereas the former only succeeded in developing productive knowledge for the controlled use of the adjectives (i.e., in the discrete item test), the latter developed both controlled and free use of them (i.e., in both the discrete item test and the Same or Different Task). Furthermore, the FonF group outperformed not just the control group but also the FonFs group in both the discrete item test and the Same or Different Task. In other words, the FonF instruction was clearly superior in enabling the learners to produce adjectives in both a controlled and a free way. However, the results for nouns did not show any difference between the two groups. There are two points in need of explanation: (1) why there was no difference in the effects of the two types of instruction for the nouns and (2) why the FonF instruction was more effective than the FonFs instruction in enabling learners to acquire the adjectives.

The results cannot be explained in terms of frequency of exposure (Palmberg, 1987) nor do they demonstrate the superiority of productive learning over receptive learning (Mondria & Wiersma, 2004). For both nouns and adjectives, both groups were exposed to similar

amounts of input but the FonFs group had more opportunities to produce both word types so this cannot explain why the FonF group performed as well as the FonFs group in the two production tests and but demonstrated greater acquisition of adjectives. To explain the results it is necessary to take a close look at the difference in the process features of the two types of instruction.

The process features of the two types of instruction were very similar in the case of the nouns (see Table 8). Both groups experienced extensive exposure to the target nouns. In both groups, the nouns were introduced pre-emptively. The learners in both groups tended to borrow the nouns from the teacher's utterances rather than initiate production of them. Although there were some differences in the process features (e.g., in the FonFs group the learners were exposed primarily to isolated use of the nouns whereas in the FonFs group they were exposed to embedded use), these do not seem to have mattered in the case of the nouns. What seems to have been important is whether the meanings of the nouns were made transparent. The FonFs instruction mostly involved requesting the learners to produce the target nouns one by one (e.g., teacher: "okay next, what's this?"). Thus, as shown in Excerpt 1, each conversational sequence focused on one vocabulary item. Also in the FonF group, each conversational sequence was centred around a teacher's command which also targeted one noun at a time.

In contrast, the process features involving the use of the adjectives were quite different in the two types of instruction. In the FonFs instruction, the adjectives, like the nouns, were introduced pre-emptively. The students were requested to produce the adjectives and they never initiated the use of them. When the learners chose to produce adjectives, they just borrowed them from the teacher's utterances. In contrast, in the FonF instruction, the adjectives were introduced reactively when the teacher used them to help the learners identify the referent of a noun in one of her commands. Thus, the adjectives were contextually embedded. Later the learners began to use adjectives themselves as in Excerpt 3. In this sequence the learners failed to understand *ostrich* and began to negotiate its meaning by using a number of adjectives which ultimately enabled them to identify the correct referent of the noun. These differences can explain the difference in the two groups' results for the adjectives. It would seem that the FonF instruction provided learners with a communicative reason for both comprehending and producing the adjectives whereas the FonFs instruction did not. In the case of the adjectives, therefore, Jiang's (2000) claim that learners need to experience the contextualized use of new L2 words in order to develop a full representation of them would seem to be borne out. In short, the FonF provided the learners with a much richer learning environment.

Another process difference concerned the opportunity for student-initiated use of the adjectives. This never occurred in the FonFs group but as Excerpt 3 shows was quite common in the FonF group. To further examine the causal effects of the FonF learners' "self-initiated" production of the adjectives on their acquisition, a simple linear regression was carried out. The number of "self-initiated" tokens produced by individual learners in this group (see Table 3) was the independent variable and the scores of the two tests the dependent variables. The results showed that in the Same or Different Task the effect was approaching significance in Post-test 1 ($R^2 = .254$, $p = .056$) and significant in post-test 2 ($R^2 = .428$, $p = .008$). In the Discrete-item test, the effects were significant in both post-

tests ($R^2 = .452$, $p = .006$. for post-test 1 and $R^2 = .445$, $p = .007$. for post-test 2). In other words, the learners who self-initiated the production of adjectives acquired more adjectives.

Another way of looking at the process differences is in terms of 'involvement load' (Laufer & Hulstijn, 2001). Table 9 evaluates the 'need,' 'search' and 'evaluation' of the use of the target words in the two groups. A minus (–) indicates an absence of an involvement factor, a plus (+) indicates that the factor is present to a moderate degree, and a double plus (++) marks the strong presence of an involvement factor (Laufer & Hulstijn, 2001, p.17).

Table 9

"Need' 'Search' and 'Evaluation' in FonFs and FonF

	FonFs	FonF	
	Nouns and adjectives	Nouns	Adjectives
Need	+ Production was necessary to complete the activities.	+ Comprehension of the nouns was necessary to complete the tasks.	+ Comprehension of the adjectives was useful to complete the tasks. + The use of the adjectives was motivated by the tasks.
Search	– The pictorial image of the words (word meaning) was provided.	+ Engaging or observing negotiation of meaning.	+ Inferring the meaning of the adjectives.
Evaluation	+ Positive and negative feedback on production was provided.	+ Choosing the correct noun card. + Feedback on the noun choices was provided.	+ Feedback on the noun choices was provided. ++ Finding the appropriate adjectives for negotiation.

Note: – : an absence of an involvement factor

Note: + : moderate presence of an involvement factor

Note: ++: strong presence of an involvement factor

Both the FonFs and the FonF required 'need', which refers to the externally imposed or self-imposed task requirement (Laufer & Hulstijn, 2001, p. 14). The FonFs required the learners to say the oral form of the words depicted on the cards. The FonF required comprehension of the meaning of the nouns and adjectives in the teacher's utterances in order to complete the tasks and generated learner-initiated production of the adjectives when negotiating the meaning of the target nouns.

The opportunity for 'search', which refers to "the attempt to find the meaning of an unknown L2 word or trying to find the L2 word form expressing a concept" (Laufer & Hulstijn, 2001, p. 14), seems to be clearly different in the two groups. In the FonFs instruction the meanings of the target words were provided in the learners' L1 at the

beginning of each lesson, and then the learners were repeatedly asked to produce the words shown on the picture cards. Therefore, the FonFs did not require any 'search' for the meanings of the target words. In contrast, the FonF instruction allowed the learners to negotiate the meaning of the nouns to complete the tasks, which served as a form of 'search' for the nouns. The learners in this group needed to infer or discover the meaning of the adjectives in order to perform the tasks successfully because the meanings of the adjectives were not directly taught.

The levels of 'evaluation' also appear to be different in the FonFs and the FonF. 'Evaluation' entails "a comparison of a given word with other words, a specific meaning of a word with its other meanings, or combining the word with other words in order to assess whether a word does or does not fit its context" (Laufer & Hulstijn, 2001, p. 14). The FonFs required 'evaluation' but only when learners responded to the teacher's positive or negative feedback on their production. The FonF required the learners to choose the appropriate noun cards from the array of picture cards, and to evaluate their task performance (i.e., whether or not they had chosen the correct card) when they received feedback. The FonF learners needed to engage in 'evaluation' because they were required to determine the meaning of an adjective when the teacher responded to their failure to comprehend by elaborating a command (e.g., T: "*the peacock is blue*") or through the negotiated interaction initiated by other students (e.g., S: "*blue?*" , T: "*yes*"). When a learner produced an adjective in the process of negotiating meaning, he/she needed to select the appropriate adjective in order to negotiate effectively. This can be considered to constitute 'strong evaluation' rather than 'moderate evaluation' (Laufer & Hulstijn, 2001, p. 15) as it required generative choice rather than selective choice.

Conclusion

This study compared the effectiveness of focus on forms (FonFs) and focus on form (FonF) by investigating both the process features of the instruction and the learning outcomes. Although both types of instruction were effective for the acquisition of nouns, the FonF instruction was found to be more effective for the acquisition of adjectives. Only the FonF learners developed the knowledge needed to use the adjectives in free production. The key differences between the process features of the FonF and FonFs instruction were proposed as an explanation for this difference in learning outcomes. That is, only the FonF instruction was characterized by (1) contextualized input, (2) the occurrence of negotiation of meaning, and (3) student-initiated production.

The results of this study are similar to those of de la Fuente (2002) but different from those of Laufer (2006). These mixed results can be explained by examining the process features of the FonF instruction in the different studies.

Table 10

Process Features of FonF Instruction in Four Studies

Process features	Current study		De la Fuente (2006)	Laufer (2006)
Target items	Nouns	Adjectives	Nouns	Nouns, verbs,

				adjectives
Contextualized input	√	√	√	√
Opportunity for negotiation of meaning	√	√	√	
Student-initiated production		√	√	

As Table 10 shows, all of the above studies involved contextualized input but only the current study and de la Fuente' study involved negotiation of meaning and student-initiated production. This analysis suggests that one aspect of FonF that is crucial for learning is student-initiated production. The teacher's contextualized input led to the learners initiating production of the adjectives in order to negotiate meaning. In de la Fuente, there was also opportunity for student-initiated production, in this case for nouns. There was no such opportunity in Laufer (2006).

Table 10 also suggests that intentional vocabulary learning is not invariably superior to incidental vocabulary learning at least for young beginner learners. In all the studies shown in Table 10 the FonFs instruction involved intentional learning and the FonF incidental learning. In Laufer (2006) the intentional vocabulary learning induced by FonFs instruction was superior. However, in this study and in de la Fuente's study the incidental learning induced by the FonF instruction was superior. What is important for vocabulary learning, then, is not whether the instruction involves intentional or incidental learning but the nature of processing involved as suggested by Laufer and Hulstijn (2001) and as demonstrated in the analysis of the process features of the FonF instruction.

It is of course possible that different results would have been obtained if PPP had been operationalized differently (e.g., by providing more explicit corrective feedback rather than recasts). Nevertheless, the PPP instruction used in this study reflected how the approach is operationalized in many classrooms. Clearly, though, there is a need for further research comparing PPP and input-based instruction with beginner-level learners.

This study also provides support for comprehension-based instruction for beginner learners. It demonstrates that instruction does not need to require production of the target items by the learners in order for productive knowledge to develop. It provides further support for Ellis' (2001) claim that listen-and-do tasks are effective in creating contexts for the acquisition of vocabulary and it extends Ellis' research by indicating which process features that arise from the implementation of such tasks are important for learning for young children.

Notes

[1] This study was a completely new study involving new data.

[2] A reviewer pointed out that presenting closely related words and antonyms together creates a greater learning burden. The decision was made to present the adjectives in the

FonFs group in this way because course books for young EFL learners typically present words in sets.

[3] Recasts were provided in response to the learners' erroneous production. Recasts were chosen as the preferred corrective feedback strategy because previous research (e.g., Seedhouse, 2001) has shown they are teachers' favoured corrective strategy for addressing errors.

[4] The study also included two comprehension-tests as well as these two production tests. However, only the results of the production tests will be reported in this article as these provide a more rigorous examination of the effects of input-based tasks on learning.

[5] The tests were audio-recorded. They were scored by the researcher listening to the recordings. To ensure reliability the researcher scored the tests twice at different times with 100% agreement.

References

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, Mass: Harvard University Press

Asher, J. J. (1977). *Learning another language through actions.* Los Gatos, CA: Sky Oaks Productions.

Brock, C. A. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly, 20*, 47-59.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). San Diego, CA: Academic Press.

Daulton, F. E. (1999). English loanwords in Japanese -- The built-in lexicon. *The Internet TESL Journal, V,* 1.

de la Fuente, M. J. (2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research, 10*(3), 263-295.

DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology.* New York: Cambridge University Press.

Doughty, C., & Williams, J. (1998). Issues and terminology. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 1-12). Cambridge: Cambridge University Press.

Ellis, R. (2001). Non-reciprocal tasks, comprehension and second language acquisition. In M. Bygate, P. Skehan. & M. Swain (Eds.), *Researching pedagogic task.* London: Longman.

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford: Oxford University Press.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly, 35*(3), 407-432.

Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus-on-form. *System, 30*(4), 419-432.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44*(3), 449-491.

Gass, S., & Torres, M. J. A. (2005). Attention when: An investigation of the ordering effect of input and interaction. *Studies in Second Language Acquisition, 27*, 1-31.

Gasser, M., & Smith, L. B. (1998). Learning nouns and adjectives: A connectionist account. *Language and Cognitive Processes, 13*(2), 269 - 306.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*(1), 47-77.

Laufer, B. (2006). Comparing focus on form and focus on formS in second-language vocabulary learning. *The Canadian Modern Language Review, 63*(1), 149-166.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research, 15*(4), 391-411.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*(2), 126-141.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. d. Bot, R. Ginsberg. & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26*(1), 27-56.

Long, M. H., & Sato, C. J. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 268-285). Rowley, MA: Newbury House.

Mackey, A., & Goo, J. M. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Input, interaction and corrective feedback in L2 learning* (pp. 379-452). Oxford: Oxford University Press.

Mondria, J.-A., & Wiersma, B. (2004). Receptive, productive, and receptive + productive L2 vocabulary learning: What difference does it make? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 79-100). Amsterdam, the Netherlands: John Benjamins Publishing Company.

Nakata, R., Frazier, K., Hoskins, B., & Graham, C. (2007). *Let's go.* London: Oxford University Press.

Palmberg, R. (1987). Patterns of vocabulary development in foreign-language learners. *Studies in Second Language Acquisition, 9*(2), 201-219.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Seedhouse, P. (2001). Case of the missing "No": The relationship between pedagogy and interaction. *Language Learning, 47*, 547–583.

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*(2), 137-158.

Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils.* London: Oxford University Press.

Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge: Cambridge University Press.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.