Journal of Culture and Information Science, 3(1), 1-13,(March 2008)

研究論文

The Necessity of Negative Feedback for Learning L2 Collocations

Tomonori Nagano and Kenji Kitao

This paper reports on a study that investigated the role of negative evidence in learning L2 collocations. Forty-four Japanese ESL/EFL learners responded to online grammatical-judgment tasks. The data suggest that (1) L2 collocations are equally challenging to all proficiency groups, (2) L1 collocations influence judgment of L2 collocations, and (3) positive evidence alone is not sufficient for learning L2 collocations.

1 Introduction

As with first language acquisition research, the role of *negative evidence* (e.g., *correction* of language learners' errors) has been one of the incentives for second language acquisition research. On one hand, researchers who are strongly influenced by innatism in first language acquisition (FLA) argue against the role of negative evidence in second language acquisition (SLA). Those researchers suggest that only *positive evidence*, along with the innate hard-wired language acquisition device, can make SLA possible. On the other hand, a number of SLA researchers propose that negative evidence is necessary, or at least facilitative, for second language learning. Those researchers claim that some forms of negative feedback (e.g., *implicit evidence* such as *recasting* and *modeling*) is logically necessary because L2 learners need to know what forms are ungrammatical in the target language. The research on negative evidence has had a considerable impact on language teaching practice in Japan. If negative evidence is necessary, the usefulness of explicit grammar instruction (which is typically conducted in the students' native language, Japanese) should be embraced. If a second language can be learnt only from positive evidence as children learn their L1, communicative teaching methods should be more effective than the grammar-based approach. In this study, we will present a study that attempts to shed light on the usefulness of negative evidence in the acquisition of collocations in a second language.

by Tomonori Nagano (CUNY Graduate Center) and Kenji Kitano (Doshisha University)

2 Literature Review

2.1 Against Negative Evidence

Schwartz (1993) explicitly denies the role of negative evidence in L2 acquisition. She distinguishes between two types of language input, *Primary Linguistic Data (PLD)* and *Negtive Data (ND)*. Roughly speaking, PLD is equivalent to what FLA researchers call positive evidence, and ND is equivalent to negative evidence. Schwartz claims that explicit L2 instruction with explicit corrections (ND) might influence learners' L2 production[1], but ND will never change learners' L2 linguistic competence. L2 competence can be developed only

[1] Schwartz also classifies learners' knowledge into *learned linguistic knowledge* (*LLK*) and *learned linguistic behavior* (*LLB*). We will not discuss the distinction in this paper, though.

with the provision of PLD in a similar fashion to how children acquire their first language. Schwartz's proposal has a strong root in the assumption in generative grammar, so-called the *poverty of stimulus*. The poverty of stimulus argues that the knowledge of grammar that children acquire is not fully instantiated in the parents' speech. For example, a native speaker of English should know "*Colorless green ideas sleep furiously*" is grammatical even if she/he hasn't heard it before and the sentence is meaningless (Chomsky, 1957). Since it is logically impossible to manifest all kinds of structures in the language input, generative linguists claim that the ability to judge grammatical sentences cannot be dependent on the language input to which the child is exposed. The generative linguists hypothesize the language acquisition device, which is innately available for all human beings and is activated only by the positive evidence (PLD in Schwartz's term.).

To support the generativists' claim, in the 1960s and 1970s, researchers investigated parents' use of *negative feedback* in response to children's language utterances. A series of studies has shown that there is no consistent relationship between parents' feedback and children's grammatical errors[2]. For example, Brown and Hanlon (1970) investigated caretakers' speech and their response to children's grammatical/ungrammatical utterances. They found that children learning an L1 were not sensitive to the correction to their grammatical errors. It was also found that there was no correlation between parents' cognitive approval (e.g., disapproval) and ungrammatical sentences of children.

To summarize, since (1) the knowledge of grammar is not fully represented in the language input that children are exposed to and (2) there is no consistent negative feedback to children's grammatical mistakes, generative linguists extrapolated an innate language acquisition mechanism, which is often called *Universal Grammar* or UG. The UG is primarily triggered by positive evidence, and negative evidence is not only unavailable in language input but also unnecessary for the acquisition of the first language. Researchers like Schwartz (1993) argue that UG is also available in SLA, and, similar to the FLA, L2 acquisition is possible with positive evidence alone since the UG restricts possible structures of the target language.

2.2 Facilitative Role of Negative Evidence

Long (1996), referring to the same logical problem[3], proposes that negative evidence is generally facilitative for L2 acquisition and is, in fact, necessary for the acquisition of specific types of L2 structures. Long argues that parents use certain implicit forms of negative feedback such as *clarification* and *corrective recast* more frequently to children's ungrammatical utterances than to grammatical ones. *Clarification* is an implicit request to repeat the utterance and *recasts* are utterances that rephrase the child's utterances by changing one or more components of the sentence while still referring to the same central meaning. Long proposes that even if there is no overt negative feedback in the parents' speech, there is implicit (but consistent) negative feedback to children's grammatical errors. He further claims that the function of implicit negative feedback is to make certain grammatical forms salient for children/L2 learners, which helps learners attend to or notice the target forms. In the revised version of his *Input Hypothesis*, Long stressed the role of noticing to L2 forms and suggested that negative evidence facilitates L2 learners to notice gaps between their interlanguage structure and the L2 structures.

In fact, there are several empirical studies that support Long's proposal. Hirsch-Pasek, Treiman, and Schneiderman (1986) initially found no correlation between parents' approval and well-formedness in child language like the study by Brown and Hanlon (1970). They found, however, tallied repetition (*verbatim* and *modification*) are more likely to be found after ungrammatical sentences than grammatical ones. Demetras, Post,

[2] For more recent developments related to negative feedback in FLA, see Marcus (1993) that addresses the potential role of *noisy input* and possible availability of *indirect negative input*.

[3] Long uses the term *Baker's paradox* (Baker, 1979), but its fundamental claim is identical to that of *the poverty of stimulus* argument.

and Snow (1986) also found that *verbatim repetitions* and *continuations* followed children's grammatical sentences slightly more often than their ungrammatical ones. In their study, *clarification of questions* was also found after ungrammatical sentences. Penner (1987) found the frequent use of corrective recast (almost twice more frequently) to ungrammatical sentences. Bohannon and Stanowics (1988) reported that 90% of *exact repetition* by parents followed children's grammatical utterances whereas 70% of *recast & expansions* followed ungrammatical ones. Finally, Farrar (1992) reported stronger relationships between caretakers' corrective recast and the degree of absence of children's ill-formed utterances. He found one-fifth of the caretakers' responses to ungrammatical utterances was *corrective recasts* and that children imitated the corrections made by corrective recasts more often than corrections made by other types of response.

While some studies seem to support Long's claim about implicit negative feedback, other researchers disagree how to interpret the data from those studies. For example, Steven Pinker, a renowned language acquisition researcher, was extremely skeptical about the role of negative evidence in language acquisition. Grimshaw and Pinker (1989) supported Lightfoot's degree-0 learnability hypothesis (1989) and argued that positive evidence would be sufficient for triggering the language acquisition mechanism in the child's mind. Grimshaw and Pinker pointed out that although many studies found statistically significant correlations between parents' implicit feedback and children's ungrammatical utterances, the statistical significance falls short of explaining the uniform and systematic process of the language acquisition. The statistical significance usually allows for a slight degree of computational errors, but in reality all normal children acquire their first language without fail. In other words, the variability of implicit feedback cannot explain how all children, regardless of what language they speak, uniformly acquire the language in similar developmental paths.

2.3 Empirical Studies about Negative Evidence

The theoretical schism prompted a number of empirical/experimental studies that focused on the influence of negative evidence on the L2 acquisition process. Lightbown and Spada (1990) claim that learners' accuracy, fluency and overall communication skills developed best with explicit negative feedback in primarily meaning-based language classrooms. Lightbown and Spada examined the classroom observation data from a five-month intensive ESL program in Quebec and found that teachers provided different frequencies of explicit negative feedback to the students. Although the primary focus of those ESL classes was on the communicative use of language, the teachers did provide some negative feedback in various manners. Lightbown and Spada showed that the learners who received more instructional intervention (e.g., grammar instruction and error correction) achieved higher accuracy in the English structures that French-speaking ESL learners particularly tend to make make mistakes on. Their study suggests that focus-on-form intervention also resulted in higher fluency and overall communication skills.

Some experimental studies show that not all negative feedback is effective. It seems that some types of negative feedback are more effective than others. Carroll and Swain (1993) were interested in the effects of various types of negative feedback and conducted a study to examine the acquisition of the dative alternation rules with 100 Spanish-speaking ESL learners. Carroll and Swain categorized negative feedback into 4 groups; that is, *explicit hypothesis rejection* (explicit metalinguistic information about the rules and generalization), *explicit utterance rejection* (explicitly notifying errors), *modeling plus implicit negative feedback* (providing corrective recasts), and *indirect metalinguistic feedback* (confirmation). They found all the subjects who had received any form of negative feedback (either explicit, implicit, or both) outperformed the control group that didn't receive any negative feedback. Since the explicit negative feedback group achieved particularly higher achievement, Carroll and Swain concluded that implicit negative feedback would not be as helpful as explicit metalinguistic feedback. In spite of several methodological flaws in their research design, their tentatively conclusion showed

that certain types of negative feedback had beneficial effects on the acquisition of the dative alternation rules at least in the short term.

Another experimental study by Tomasello and Haeron (1988) also reported benefits of negative feedback in contrast with explicit instruction about the exceptions to generalizable rules. Tomasello and Haeron provided two groups of college-level French learners with (1) explicit presentation of target structures and exceptions to the canonical rules and (2) only canonical exemplars. Tomaesello and Haeron called the latter treatment *the Garden Path condition* because it encourages students to make mistakes and provide corrective recast instead of explicit presentation of rules. They reported the Garden Path group surpassed the explicit rule group, suggesting beneficial role of negative feedback over explicit presentation of rules. The Garden Path condition, according to them, seems to have helped both in drawing students' attention to the rule itself and making them recognize exceptions to the grammar rules. Tomasello and Herron (1989) succeeded in replicating the same pattern in their follow-up study. They claim that *corrective recasts* are especially useful for language learners to compare their own speech with native-speakers' models since they occur (i) immediately after the child's incorrect utterance, (ii) in the same discourse context, and (iii) with the same semantic context.

2.4 Studies that Focus on Japanese-English Language Interaction

There are a few studies that specifically focus on Japanese-English language interaction. Hirakawa (2001) has shown that an expected learnability problem in the classification of unaccusative/unergative verbs in L2 is not attested in her experiment with Japanese language learners. Using the adverbial modifier （たくさん） and the aspectual morpheme ている as the diagnostics for verb accusativity, Hirakawa tested 25 adult English-speaking learners of Japanese and 20 Japanese-native speakers. Hirakawa concluded that her subjects successfully distinguished unaccustative verbs and unergative verbs in the target languages. She did not discuss, however, how they have conquered the learnability problem of unaccusative/unergative verbs. It should be noted that the distinction between unaccusative/unergative verbs is usually not explicitly taught in the formal language instruction; thus, there should be little negative evidence for language learners.

Unlike Hirakawa (2001), Izumi (1998), who investigated the acquisition of passive constructions in English and Japanese, claims that negative evidence is required for the acquisition of structures that are grammatical in L1, but not in L2. While English allows only *direct passive*, Japanese have two different types of passive constructions, *direct passive* and *indirect passive*. The Japanese direct passive has the identical structure to English direct passive, which is formed by suppressing the external or AGENT argument of the transitive verb. On the other hand, the Japanese indirect passives are formed with both transitive and intransitive verbs. Izumi (1998) argues that Japanese ESL learners will have problem in the passive construction in English because they might treat auxiliary verb *be* as a lexical verb (as in Japanese indirect passive). He hypothesized the necessity of *negative evidence* if the possible grammatical structures in L1 is a superset of L2. In other words, Japanese ESL learners need explicit instruction about the ungrammaticality of the indirect passive construction in English. In order to test this hypothesis, Izumi used three different tests (i.e., translation, picture-cued production, and grammaticality judgement) at three different stages (i.e., pre-treatment, immediate post-treatment, 2nd post-treatment after 8 weeks) with 15 Japanese ESL students. Four students received explicit instruction about the passive constructions in English (negative evidence) and 11 received normal ESL instructions. In spite of some questionable use of statistics, Izumi has concluded that his hypothesis is supported. He argues that negative evidence is necessary if L2 is not a subset of L1 with respect to a certain grammatical construction.

Inagaki (2001) has reported an opposite case of Izumi (1998) in which the grammatical constructions in Japanese are a subset of corresponding constructions in English. Inagaki shows that the

English construction [Manner-of-motion verb + GOALPP] is not available in Japanese while the construction [Directed-motion verb + GOALPP] is allowed in both English and Japanese[4]. Since the Japanese grammatical structures are a subset of the corresponding English grammatical structures, Japanese ESL learners should *not* have difficulty recognizing the grammaticality of English [manner-of-motion verb + GOALPP] constructions. On the other hand, English-speaking Japanese learners will have difficulty detecting that [manner-of-motion verbs + GOALPP] is ungrammatical in Japanese. To test this hypothesis, Inagaki tested 64 English-speaking Japanese learners and Japanese-native speakers with the grammaticality judgement task in both English and Japanese. The results show that Japanese speakers accepted almost all English sentence types[5]. In the Japanese-sentence experiment, as hypothesized, English speakers accepted all sentence types, failing to detect the ungrammaticality of [PP + ManV] in Japanese. Thus, the study suggests that the second language learners have difficulty with structures when their first language accepts the wider range of structures.

2.5 Factors for the Role of Negative Evidence

The studies about the role of negative evidence in SLA exhibit very different results, sometimes even contrastive. We suppose that the mixed results of the past literature may be attributed to a few theoretical perspectives that each researcher takes. Although there is a strong theoretical appeal to considering the fundamental process of second language acquisition is equivalent to that of FLA, there are a few obvious differences between them. The first factor to be considered is *initial state*. Unlike the first language learners, the second language learners start off the language acquisition with some knowledge about language from their L1. In other words, the *initial state* of SLA may well be different from that of FLA. Assuming that the *subset principle* (Wexler & Manzini, 1987) is correct, if a second language learner initiates language acquisition with grammar that accepts a wider range of parameters (that is, if the L1 grammar is a superset of the L1 grammar), the negative evidence is necessary to reshape the learner's grammar into an appropriate L2 grammar, which is a subset of his/her initial grammar. On the other hand, if one assumes the initial state of SLA has little to do with L1, the necessity of negative evidence is alleviated. The second factor is *proficiency*. Unlike the first language learners, second language learners exhibit a huge variety of language proficiency. Few studies (with exception of Iwashita [2003] took into account the possible influence of proficiency on the effectiveness of negative evidence. In her study with English-speaking adult Japanese learners, Iwashita shows that the effect of positive evidence end negative evidence interacts with the proficiency levels of the learners. More precisely, the positive evidence benefited only high-proficient learners whereas the implicit negative input was beneficial for all learners across their proficiency levels. This study indicates some interaction effects between the types of positive/negative evidence and the proficiency level.

In our study, we consider those two factors crucial and they are carefully controlled.

3 This Study

3.1 Motivation for the Study

In this study, we will address whether the role of negative feedback facilitates the acquisition of a specific type of language use, *collocations*. We believe that the *collocation* is an interesting area to test negative evidence for three reasons. First, although it is not widely recognized, the correct use of collocations in the second language is one of the most daunting challenges in attaining a native-like L2 fluency. In both generative linguistics and language education, collocations have been regarded as a peripheral phenomenon of language acquisition. For example, the obvious unacceptability of s*trong computer* (as opposed to the acceptable variant, *powerful computer*) has attracted very little attention in either field. However, the apt use of collocations

[4] Due to the head-final nature of Japanese, the corresponding structure of [Manner-of-motion verb + GOALPP] in English is [PP + Manner-of-motion verb] such as *ジョンが学校に行く.

[5] Except for the [DirV + PP + -ing] structure.

is a hallmark of native language fluency and the learning mechanism of collocations should be worth more attention in L2 acquisition research. Second, the fact that collocations are rarely taught and are hardly corrected in second-language classroom presents a unique situation in terms of positive/ negative feedback. An example will suffice to illustrate this point. Suppose that an ESL learner uttered an unconventional collocation *a strong computer* based on an erroneous application that *strong* can be a modifier of *computer*. Since it is easy to infer the intended meaning of the phrase, such collocation mistakes are often overlooked without any explicit correction. Obviously, collocation errors are less likely to be subject to explicit feedback than structural errors (e.g., 3rd person singular -*s*). We assume that, in spoken communication, collocation errors are probably not corrected at all. In this study, we are using the somewhat naive assumption that negative feedback (either explicit or implicit) to inapt collocations is totally absent in SLA. Finally, in collocation research, it is possible to make a clear distinction where the influence of the first language is expected and where such an influence will not occur. The influence of the first language, or *language transfer*, is one of the central problems in SLA. It is often assumed that L1 influences second language acquisition to some degree, but there has been no evidence for the extent to which and on what aspects of language the L1 has an influence. In order to control the L1 influence, we assume that collocations are the result of the compositionality of lexical semantics. For example, while *strong* and *powerful* are synonymous, only the semantics of *powerful* is extended to the specification of machinery (thus, *powerful computer* and *powerful car* are possible) and only the semantics of *strong* is extended to the degree of stimulus (thus, *strong medicine* and *strong beer* are possible). Since the compositionality of lexical semantics is language-specific, we can distinguish collocations depending on whether the same semantic extension is observed in L1 or not (we will discuss this distinction in detail in the methods section).

3.2 Research Question

The findings in the experimental studies discussed above suggest that negative evidence is necessary or at least helpful to acquire a second language. Based on the findings of those studies, in this study, we will assume that negative evidence is also necessary or facilitative for acquiring L2 collocations. Since this study is a preliminary one, we pose the following research question.

- Are collocations that are ungrammatical in Japanese (L1) but grammatical in English (L2) more difficult to learn than collocations that are grammatical in both languages?

It is our assumption that if negative evidence is necessary or facilitative in learning L2 collocations, the answer will be positive, because no negative evidence is available to the learners in learning collocations. For example, お使いに行く *otsukai-ni iku* (*go on an errand*) is grammatical both in Japanese and English, but お使いに走る *otsukai-ni hashiru* (*run an errand*) is ungrammatical in Japanese. In contrast with *go on an errand, run an errand* should be difficult since, for ESL learners of Japanese native speakers, the only way to learn such collocations is through the provision of positive evidence in the language input.

4 Methods

4.1 Participants

Forty-one Japanese learners of English participated in this experiment. Almost all participants were college students or faculty who lived in Japan at the time of testing, but some participants were university/ graduate students in the U.S. In order to control the effects of English proficiency level, the participants were divided into three proficiency groups based on the questionnaire answers, which includes the results of English proficiency exams (e.g., *Eiken*, TOEFL, and TOEIC) and their self-evaluation of their English proficiency. Thirteen participants were classified as beginners, 22 participants as intermediate, and 6 as advanced.

4.2 Materials

The study was carried out online by using a

module written in Flash. (See the screen captures in Appendix A and Appendix B). In the module, collocations were presented one by one on the screen along with a picture that showed the interpretation of the collocation. The online module is available at www.slacorpus.com/programs/collocation.html . The subjects rated the acceptability of each collocation stimuli on a 5-point Likert scale, ranging from 良い (good) to 良くない (not good). The choices of the grammatical judgment and reaction times were recorded. Before responding to the collocations, the participants answered a questionnaire, based on which they were classified into three different proficiency groups.

4.3 Stimuli

Participants were presented with 44 English collocations. English collocations were randomly selected from Benson, Benson, and Ilsen (1986) and Japanese collocations are generated based on the first author's judgment. Samples of the collocations are presented in Table 1. The complete list of collocations is in Appendix C.

Table 1: Examples of collocation stimuli

Type-0
(grammatical collocations in both languages)
a bicycle chain
a change of direction
an excellent student
Type-1
(grammatical in English, but not in Japanese)
a strong student
admit defeat
an uphill battle
Type-2
(not acceptable in either language)
a loose change
a strict battle
an elite student

Based on the acceptability of collocations in L1 and L2, the stimuli were grouped into three types; Type-0 is a set of common English collocations whose literal translations are also common collocations in Japanese, Type-1 consists of common English collocation whose literal translations are not acceptable collocations in Japanese, and finally, Type-2 are collocations that are not acceptable in either English or Japanese. The collocation category is listed in Table 1.

Since the concepts of the Type-0 collocations are also available in the participants' first language (Japanese) and the concept can possibly be transfered to the second language (English) collocations, we assume that Type-0 is the easiest collocations to learn. By the same token, we assume that Type-2 collocations are the easiest to reject because the concepts of Type-2 collocations do not exist in both L1 and L2. Therefore, in this study, Type-0 and Type-2 collocations are considered the baselines for the easiest to learn and the easiest to reject, respectively.

Type-1 collocations are the focus of our interest in this study. Since the concepts of Type-1 collocations are not acceptable in the native language (Japanese) and there is little negative feedback for learning collocations in general, we assume that the only means to acquire Type-1 collocations is through positive evidence. Thus, if negative evidence is necessary for acquiring collocations, Type-1 collocations should be mistakingly judged as unacceptable (that is, the judgment of Type-1 is similar to Type-0). If collocations are learnable only with *positive evidence*, the learners should be able to judge them as grammatical (that is, Type-1 is similar to Type-2).

5 Results

The descriptive statistics of grammatical judgment scores and reaction times are presented in Table 2. The data were submitted to the 3×3 ANOVA in order to find the effects of the item types and the participants' L2 proficiency levels. Repeated *t*-tests were conducted as a follow-up analysis to test the item type effect.

As is seen in Table 2, there is a tendency in all proficiency groups to judge Type-0 as the most acceptable, followed by Type-1 and Type-2, which is the least acceptable. A similar tendency is found in the reaction time. Type-0 collocations required a longer response time than Type-1, and Type-1 collocations required longer than Type-2. Only

Table 2: Mean and SD (in parentheses) of the grammatical judgment scores (1="good"to 5="not good") and RT (in seconds)

	all (N=41)	**beginner** (N=13)	**intermed.** (N=22)	**advanced** (N=6)
Type-0 GramJ	1.50 (0.54)	1.52 (0.53)	1.47 (0.60)	1.50 (0.54)
Type-1 GramJ	1.66 (0.48)	1.67 (0.50)	1.62 (0.43)	1.64 (0.56)
Type-2 GramJ	1.76 (0.46)	1.80 (0.42)	1.70 (0.48)	1.74 (0.64)
Type-0 RT	5.71 (2.17)	5.87 (2.70)	5.51 (1.24)	5.59 (1.79)
Type-1 RT	5.77 (2.41)	6.01 (2.84)	5.96 (1.67)	5.83 (2.41)
Type-2 RT	6.02 (2.19)	6.13 (2.57)	6.06 (1.62)	5.56 (1.99)

exception is that the advanced proficiency group responded faster in Type-2 items than other items.

Both the grammatical judgment and the reaction time data were submitted to 3×3 ANOVA, using "grammatical judgment score" or "reaction time" as dependent variables, "question type" as the within-subject factor, and "proficiency groups" as the between-subject factor. The results are shown in Tables 3 (grammatical judgment) and 4 (reaction time). The interactive figure for the two ANOVA analyses are also shown in Figure 1.

Table 3: Grammatical judgment scores: 3 (proficiency groups) $\times$ 3 (question types) two-way ANOVA

	df	SS	Mean Sq	F-val	p-val
profGp	2	0.12	0.06	0.24	0.79
questType	2	1.40	0.70	2.74	0.07
interaction	4	0.02	0.0043	0.02	0.99
residuals	114	29.18	0.26		

As seen in Table 3 (the grammatical judgment scores), no significant main effects or interaction effect were observed. However, there is a marginally significant main effect in the question type in the grammatical judgment scores.

Table 4: Reaction time: 3 (proficiency groups) $\times$ 3 (question types) two-way ANOVA

	df	SS	Mean Sq	F-val	p-val
profGp	2	1.84	0.92	0.17	0.84
questType	2	2.27	1.13	0.21	0.81
interaction	4	1.04	0.26	0.05	0.99
residuals	114	611.02	5.36		

Similarly, no significant effects were observed in the reaction time (Table 4). In addition, the main effect of the question type was not observed at all in the reaction time.

From the results above and the interaction figure in Figure 1, we concluded that there is no effect of the proficiency group on the grammatical judgment scores and the reaction time. From the marginal effects of the question type, we suspected that the effect of the question type exists, regardless of the proficiency level of the participants. This point is discussed further in the following section.

In order to confirm the question type effects, a follow-up analysis was conducted by combining all three groups into one single group. The data of the combined group were analyzed with repeated paired-samples t-tests, using pairs of the question types as the repeated treatments.

The results of the repeated t-tests are presented in Table 5. As predicted, there were significant differences between Type-0-Type-1 and between Type-0-Type-2. There was also a marginally significant difference between Type-1 and Type-2 items. Although the difference was very small, the

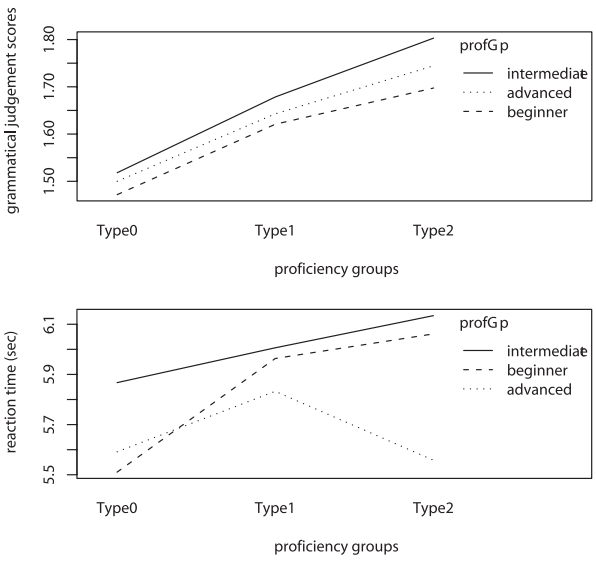


Figure 1: Grammatical judgement scores and RTs by proficiency groups

Table 5: Grammatical judgment scores: repeated t-tests

REPEATED TREATMENTS	t-value	df	p-value
Type-0-Type-1	-2.45	40	0.02
Type-1-Type-2	-1.90	40	0.07
Type-0-Type-2	-4.49	40	<0.01

participants seemed to have responded to Type-0 slightly differently to Type-1 and Type-2.

The difference between Type-0 and the other two became more obvious in the reaction time. As shown in Table 6, Type-0 and Type-1 showed a marginally significant difference while Type-1 and Type-2 failed to show any significant differences.

Table 6: Reaction time: repeated t-tests

REPEATED TREATMENTS	t-value	df	p-value
Type-0-Type-1	-1.61	40	0.11
Type-1-Type-2	-0.34	40	0.74
Type-0-Type-2	-1.89	40	0.07

The results of the repeated *t*-tests with the grammatical judgment scores and the reaction time seem to suggest that the participants responded to Type-0 items differently from the items from the other two groups. In other words, the participants failed to recognize the grammaticality of Type-1 items with the same accuracy as the Type-0 items.

6 Discussion

In this study, the role of negative evidence in the learning of collocations was investigated. Because of the lack of explicit correction and instruction on collocations, it was assumed that only positive evidence was available for second language learners to learn collocations.

In order to control for L2 proficiency effects, the participants were divided into three groups. The collocations are also divided into three groups as well in order to control for the effects of the participants' L1. In grouping the collocations, we assumed that the collocations are easy to accept (or reject) if the same lexical semantic extension exists in L1.

6.1 Proficiency Effects

One of the interesting results we observed is that the collocations remain difficult even for advanced learners (although, with such a small sample, it is difficult to draw firm conclusions). Participants in the advanced group were highly proficient L2 learners who included graduate students in the U.S. or university professors in Japan. Those participants had very high scores on standardized English proficiency exams[6] and some of them indicated they consider themselves as "near-native" speakers of English.

However, the advanced participants did not perform any better than the other participants, and there was no significant main effect for the proficiency groups in two-way ANOVA. In fact, as seen in Figure 1, their pattern of responses to the different types of collocations was nearly identical to those of the other two groups. Thus, we conclude that collocations are difficult for all second language learners, regardless of their proficiency level.

It is interesting that the advanced participants responded to all three types of collocations more conservatively than the participants in the intermediate group. The advanced participants (inaccurately) rated the acceptable (Type-0 and Type-1) collocations lower and (correctly) rated the unacceptable (Type-2) collocations lower. We interpret this conservative response by the advanced participants as a sign of their awareness of the collocations. We believe that the advanced L2 learners are consciously or unconsciously aware that some collocations are better than others, but they do not know exactly what collocations are acceptable in L2. Therefore, the advanced learners tended to make conservative judgment on all collocations across the board.

Language awareness (or *noticing*) in second language acquisition has been well documented in SLA research (Schmidt, 1990, 1994; Hulstijn & Schmidt, 1994; Izumi, 2003). We believe that the awareness of collocations by advanced speakers of L2 is qualitatively different from lower-proficient L2 speakers and will leave this topic for future research.

[6] The criteria used for this group is that the participant has passed/scored either the *Eiken* 1st grade, more than 250 on *TOEFL*, or more than 900 on *TOEIC*.

6.2 L1 Transfer

In this study, we attempted to abstract away the effect of L1 transfer by imposing a simple assumption that the lexical semantic extensions fully transfer to the L2 lexicon. In spite of its ungrounded assumption, the data suggest that L1 collocations do influence judgments of L2 collocations. The results (cf. Table 2) clearly suggest that collocations whose equivalents are grammatical in L1 are rated higher than those whose equivalents are ungrammatical in L1.

However, this is not to say that all aspects of L1 will transfer to L2. A number of studies have shown that language transfer will not occur in certain aspects of language. For example, Vainikka and Young-Scholten (1996) propose that only lexical items are transferred to L2, but not the functional projections of L1. Similarly, researchers working on the no-transfer hypotheses (Epstein, Flynn, & Martohardjono, 1996; Flynn & Martohardjono, 1994) claim that the language acquisition mechanism (UG) will not transfer at all.

Although our study indicated that the knowledge of L1 collocations transfers to L2, more careful research will necessary to draw conclusions about the nature of transfer in the acquisition of L2 collocations.

6.3 Negative Evidence

Finally, with respect to the role of negative evidence in the acquisition of L2 collocations, the data suggest that our participants failed to acquire the target (Type-1) collocations. As shown in Table 5 and 6, the participants failed to identify the Type-1 collocations as grammatical at the same rate as the Type-0 collocations.

Two interpretations are possible for this result. The first interpretation is that the participants rated Type-1 collocations lower because of language transfer. In this interpretation, the participants knew that the Type-1 collocations were grammatical in English, but they rated them lower than the Type-0 collocations because their knowledge of L1 collocations influenced the judgment. The second interpretation, which we will propose to be correct, is that the participants failed to acquire the Type-1 collocations since there was no negative feedback in their environment. In this interpretation, the participants acquired the Type-0 collocations by means of transfer from L1 collocations, but they had to learn Type-1 collocations that do not exist in their native language. The acquisition failed, however, as indicated by the significant difference between Type-0 and Type-1 collocations in Tables 5 and 6. The collocations remain problematic to the participants of all proficiency levels because, unlike other grammatical structures, no explicit or implicit negative feedback is available for collocations.

Although the both interpretations are equally plausible, we believe that the latter interpretation is right. We think that the grammatical judgment of the Type-2 collocations supports our claim. The Type-2 collocations were considered as the easiest to reject since they are ungrammatical in both L1 and L2. Although this type of collocations were rated as the lowest among three types ($\bar{X}_{Type\text{-}2} = 1.76$ as opposed to $\bar{X}_{Type\text{-}0} = 1.50$ and $\bar{X}_{Type\text{-}1} = 1.66$), the rating was still far from 5, the worst possible rating in this study. Since it is difficult to attribute the high rating of the Type-2 collocations to language transfer (i.e., the Type-2 collocations are ungrammatical in both L1 and L2), it seems reasonable that the participants simply failed to distinguish good collocations from ill-formed ones.

Due to the time and resource restrictions, our study is limited in many respects. However, if our interpretation is correct, it seems that we can conclude that the participants failed to acquire the collocations. We argue that the failure of the acquisition of collocations can be attributed to the lack of the negative feedback since the problem of collocations has not improved among the high proficiency participants. The high proficiency participants keep having problems with collocations because they cannot benefit from negative feedback, which is available in learning other grammatical constructions.

7 Conclusion

In this article, we attempted to shed a new light on the acquisition of collocations and the role of negative feedback in SLA. In spite of many

limitations, the study provided an obvious indication that the collocations remain to be a huge challenge for L2 learners. We claim that the lack of negative evidence in learning collocations is a critical factor of L2 learners difficulty with collocations and conclude this article with a remark that a more careful experimental study can be carried out in the future.

References

Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry*, **10**, 533–581.

Benson, M., Benson, E., & Ilson, R. (1986). *The BBI combinatory dictionary of English: A Guide to word combinations*. Amsterdam, Netherlands: John Benjamins.

Bohannon, J. N. & Stanowicz, L. (1988). The issue of negative evidence: Adult responses to children's language errors. *Developmental Psychology*, **24** (5), 684–689.

Brown, R. & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. Hayes (Ed.), *Cognition and the Development of Language*, 11–54. New York, NY: Wiley.

Carroll, S. E. & Swain, M. (1993). Explicit and implicit negative feedback. *Studies in Second Language Acquisition*, **15**, 357-386.

Chomsky, N. (1957). *Syntactic structures*. Janua linguarum. Series minor ; 4. The Hague, Paris: Mouton.

Demetras, M., Post, K., & Snow, C. (1986). Feedback to first language learners: The role of repetitive and clarification request. *Journal of Child Language*, **13**, 275–292.

Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Science*, **19** (4), 677–758.

Farrar, M. J. (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology*, **28** (1), 90–98.

Flynn, S. & Martohardjono, G. (1994). Mapping from the initial state to the final state: The separation of universal principles and language specific properties. In B.Lust, J.Whitman, & J. Kornfilt (Eds.), *Syntactic theory and first language acquisition: Crosslinguistic perspectives: vol.1 Phrase Structure*. Hillsdale, New Jersey: Erlbaum.

Grimshaw, J. & Pinker, S. (1989). Positive and negative evidence in language acquisition. *Behavioral and Brain Science*, **12**, 341–342.

Hirakawa, M. (2001). L2 Acquisition of Japanese unaccusative verbs. *Studies in Second Language Acquisition*, **23**, 221–245.

Hirsch-Pasek, K., Treiman, R., & Schneiderman, M. (1986). Brown and Hanlon Revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language*, **11**, 81–88.

Hulstijn, J. H. & Schmidt, R. (1994). Consciousness in second language learning. *AILA Review*, **11**, 5–10.

Inagaki, S. (2001). Motion verbs with goal PPs in the L2 acquisition of English and Japanese. *Studies in Second Language Acquisition*, **23**, 153–170.

Iwashita, N. (2003). Negative feedback and positive evidence in task-based interaction: differential affects on L2 development. *Studies in Second Language Acquisition*, **25** (1), 1–36.

Izumi, S. (2003). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition*, **24**, 541–577.

Izumi, S. & Lakshmanan, U. (1998). Learnability, negative evidence and the L2 acquisition of the English passive. *Second Language Research*, **14** (1), 62–101.

Lightbown, P. & Spada, N. (1990). Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition*, **12**, 429–448.

Lightfoot, D. (1989). The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Science*, **12** (2), 321–334.

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*. Malden, Mass.: Academic Press.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition*, **46**, 53–85.

Penner, S. (1987). Parental responses to grammatical and ungrammatical child utterances. *Child Development*, **58**, 376–384.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied linguistics*, **11**, 129–158.

Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages*, 165–209. London: Academic Press.

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, **15**, 147–163.

Tomasello, M. & Herron, C. (1989). Feedback for

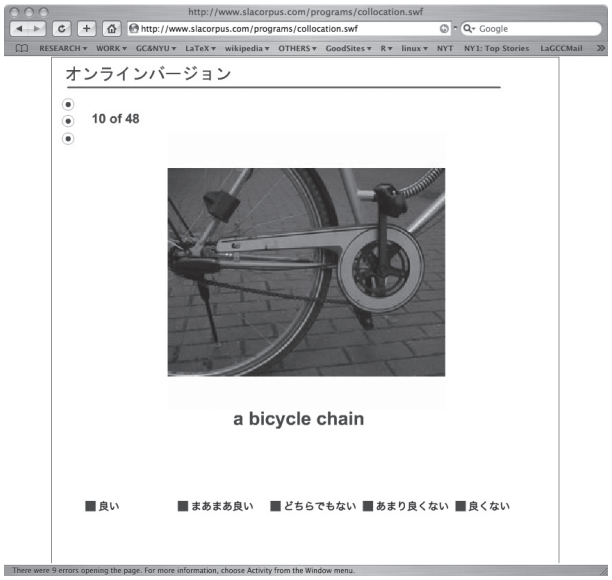
language transfer errors. The garden path technique. *Studies in Second Language Acquisition*, **11**, 385–395.

Tomasello, T. & Herron, C. (1988). Down the garden path: Including and correcting over-generalization errors in the foreign language classroom. *Applied Psycholinguistics*, **9**, 237–246.

Wexler, K. & Manzini, R. (1987). Parameter and learnability in binding theory. In T. Roeper & E. Williams (Eds.), *Parameter Setting*, 41–76. Dorderecht, Holland: Reidel.

Appendix A: The online questionnaire screenshot

Appendix B: The online collocation evaluation screenshot



Appendix C: The list of tested collocations

Type-0 (grammatical collocations in both languages)	
a bicycle chain	a change of direction
an excellent student	collect evidence
cover a distance	fight a battle
food market	humiliating defeat
meet a demand	natural beauty
positive recommendation	raise one's armes
strong rain	warm bath
withdraw an objection	
Type-1 (grammatical in English, but not acceptable in Japanese)	
a strong student	admit defeat
an uphill battle	farmers market
give the dog a bath	greet somebody with open arms
heavy rain	make a demand
produce evidence	propose a change
raise an objection	striking beauty
strong recommendation	walk a distance
Type-2 (not acceptable in either language)	
a loose change	a strict battle
an elite student	connected to a chain
decisive defeat	do an objection
drink a demand	enormous rain
enter a bath	foodstuff market
hot recommendation	lighten one's arms
put a distance	ultimate beauty
unmoving evidence	