In this study, I analyzed the collocations used by learners of Japanese in the framework of the process of second language acquisition. I used a large set of written materials from Japanese language students at a university in Taiwan, whose language of education was Chinese. Until now, in collocation research on the acquisition of English and Japanese, has defined “collocation” based on the production data of native speakers. And when a learner’s collocation use has been found to be different from that of a native speaker, researchers have considered the learner’s use an “error.” Such considerations are likely inadequate analyzing the process of second language acquisition because the focus of precious research on the development of second languages has been too narrow. In contrast, in this study, I aimed to provide insight into SLA processes with close attention to word combinations as collocations, a kind of “chunk.” “Chunking” is a human cognitive process combining associated pieces of
I determined the learners’ forms of collocation were not errors but rather productions based on the learners’ hypotheses regarding Japanese word combinations. Using comparative interlanguage analysis (CIA: Granger 1998b), I analyzed which words the learners and native speakers used with the verb “suru” (do). The analysis examined three forms of collocation: (a) lexical collocation, (b) grammatical collocation, and (c) semantic collocation.

The data set used in this study was the “Learner’s Language Corpus of Japanese” (Global COE Program, Tokyo University of Foreign Studies ‘Corpus-based Linguistics and Language Education’ (CbLLE)). This corpus was comprised of composition data from various learners and native speakers- I used the Taiwanese learners’ and native speakers’ data. The subjects of the present study were 102 students majoring in Japanese at a university in Taiwan. The Japanese native speakers used for comparison with the Taiwanese learners were 59 Japanese university students in Tokyo. The subjects each completed 8 kinds of written functional tasks and 8 written diary entries. The native speakers likewise completed the 8 functional tasks and only one diary entry. The total number of words analyzed for the subjects was 169,432, and the total for the native speakers was 47,979. After eliminating the personal information from the collected data, the data was converted to text files and part-of-speech tags were attached. In my analysis, I extracted all the instances of the verb “suru” from the data set, entered them with co-occurrence words into Microsoft Excel. The total number instances of “suru” analyzed for the subjects was 3,816, and the total for the native speakers was 1,177.

The results suggest, the collocations of “suru” used by the subjects were likely different in their core meanings and the derivative meanings of “suru.” First, most
subjects formed strong connections between “suru” and several specific words, both for the core meaning (daily actions) and the derivative meaning close to the core meaning (actions and changes) of “suru,” for example, “shigoto wo suru (to work),” “benkyoo suru (to study.)” At the same time, the subjects used two grammatical combinations frequently: “noun + wo suru” and “noun + suru.” They used a variety of words with “suru” in those forms. Also, they creatively connected various semantic categories of words with “suru.” These results indicate that, with both the core meaning and the derivative meaning close to the core-meaning of “suru,” the learners were likely to have fixed lexical collocations such as “shigoto wo suru” or “benkyoo suru” and to produce creative combinations based on the rules of grammatical collocations; “noun + wo suru” and “noun + suru,” and semantic collocations; “action.” I designated such collocations of the core meaning as “core meaning collocations.” On the other hand, for two derivative meanings, the subjects formed grammatical collocations: “adjective + suru” to mean change (e.g. akaku suru, to make something red) and “noun + ni suru” to mean decide (e.g. iku koto ni suru, to decide to go). However, the subjects had only fixed lexical collocations for most of these meanings. For example, the subjects didn’t use the grammatical collocation “o + verb + suru” (e.g. o kari suru, the polite form to borrow) like the native speakers did. The most common use of this form by the subject was “o negai suru,” which was not considered to produce based on the grammatical collocation, but rather based on the lexical collocation of “o negai” and “suru.” These results demonstrate that the subjects were more likely to use fixed co-occurrences for the derivative meanings rather than produce language creatively, as they did for the core meaning. I designated such collocations of the derivative meanings as “derivative meaning collocations.”
A lexical collocation is a co-occurrence of a word and a word, is fixed chunks of language that can be used as a whole apart from grammatical collocation and semantic collocation. In contrast, grammatical collocations and semantic collocations are co-occurrences of words and groups of words, functioning under abstract rules. Analysis of the three collocation types for the core meaning collocations and the derivative meaning collocations indicate that the acquisition of the collocations of “suru” is likely to proceed from lexical collocation to grammatical collocation and semantic collocation, as well as from the core meaning to the derivative meanings.

Considering these results, the subjects likely produced collocations of the verb “suru” based on some chunks or grammatical and semantic rules. Granger (1998b) analyzed collocations of amplifiers used by learners of English compared to the collocations used by native speakers. He said the subjects used “amplifiers more as building bricks than as part of prefabricated sections,” that is, the subjects co-occurred words at random with amplifiers. However, in this study I analyzed the learners’ uses of collocations of “suru” based on the perspective that collocations were considered as a kind of chunk, the results likely contrast with Granger (1998b), the subjects were likely to use “suru” as part of prefabricated patterns. Moreover they were likely to have systems for collocations of “suru,” as “core meaning collocations” and “derivative meaning collocations,” that are different from those of native speakers.