AN ANALYSIS OF THE KNOWLEDGE AND USE OF ENGLISH COLLOCATIONS BY FRENCH AND JAPANESE LEARNERS

AN ANALYSIS OF THE KNOWLEDGE AND USE OF ENGLISH COLLOCATIONS BY FRENCH AND JAPANESE LEARNERS

Shino KUROSAKI

An Analysis of the Knowledge and Use of English Collocations
by French and Japanese Learners

Cover image by Galdzer/Bigstock.com

ABSTRACT

While it has been recognized that the use of collocations is significant for L2 learners, much research has not been carried out on the knowledge and use of learner's collocations. The present study investigated differences on the knowledge and use of collocations between French and Japanese learners with regard to: 1) L1 influence; and 2) combinability and transparency influence.

The test materials included four categories of the lexical collocations: 1) verb + noun; 2) delexicalised verb + noun; 3) adjective + noun; and 4) adverb + adjective. The two types of tasks, Multiple Choice Question Tasks and Translation Tasks, were performed, and the learner corpora were also investigated in order to examine whether the learners from different L1 backgrounds demonstrate different results.

Since both French and English belong to Indo-European background languages, they share a number of cognate words. Thus, originally it was expected that L1 influence of the French learners would be higher in all of the four lexical collocations than that of Japanese learners, who have non-Indo-European backgrounds. Though L1 influence by both French and Japanese learners was demonstrated, the Japanese learners showed a greater L1 influence in the [adjective + noun] category than the French learners. The investigation also found that L1 influence does not necessarily result in accuracy of the collocations.

With regard to the combinability and transparency influence, the results of the two types of tasks followed the previous remark made by Kellerman (1978) who argues that L2 learners are unable to transfer words with figurative meaning. However, some contrasted results were also identified in learner corpus investigation. Thus the combinability and transparency influence were not necessarily identified.

The results of the present study have a potential to improve teaching/learning of collocations through recognizing the learners' tendencies of learning collocations.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY OF TERMS AND ABBREVIATIONS

BNC	British National Corpus
CEFR	Common European Framework of Reference
Collocation	A type of two- to three- word combination in a certain grammatical pattern, such as "verb+noun", which occurs frequently.
EAP	English for Academic Purposes
EFL	English as a Foreign Language
ESL	English as a Second Language
Formulaic Sequences	A sequence of words which is prefabricated, which contains from a few word sequences to a longer ones.
HKUST	Hong Kong University of Science and Technology
ICLE	International Corpus of Learner English
MCQ	Multiple Choice Questions
NNS	Non Native Speaker
NP	Noun Phrase
NS	Native Speakers
RC	Restricted Collocation
SLA	Second Language Acquisition
SPSS	Statistical Package for Social Science
ULIP	University of London Institute in Paris

ACKNOWLEDGMENTS

Chapter 1

INTRODUCTION

1.1 Background to the Study

Over the past three decades, more and more researchers (Krashen & Scarcella 1978; Nattinger & DeCarrico 1992; Schmitt & Carter 2004; Wray 2002; Read & Nation 2004) have focused on the importance of multiword unit of language such as formulaic sequences, idioms, and collocations in L2 learning. Multiword unit have been considered as widespread formula by L2 learners (e.g. Fillmore 1976), and its importance that they are stored and processed holistically has been discussed (e.g. Wray 2002). The empirical research has been carried out, for example, on the processing of multiword unit. Jiang & Nekrasova (2007) found that formulaic sequences are processed more quickly and more accurately than non-formulaic sequences. In recent research on written language, researchers use terms such as clusters, chunks or bundles to refer to formulaic sequences common in written texts. One of types of formulaic sequences, "lexical bundles", which are one type of "multiword units that occur most commonly in a given register" (Biber & Barbieri 2007) have recently been studied, because they are important building blocks of discourse in spoken and written registers and prevalent in university classroom teaching. They found that lexical bundles are more prevalent in non-academic university registers than in the instructional registers and more common in written course management than in spoken university registers. The studies on the use of formulaic language in speech were carried out (De Cock *et al.* 1998; Adolphs and

Durow 2004; Shin & Nation 2008) as well as those in writing (Granger 1998; Nesselhauf 2005; Webb & Kagimoto 2009). With regard to the written language, according to one of the previous studies (Hyland 2008), the frequent formulaic sequences used in the creation of academic discourse differ between disciplines.

The studies on another types of formulaic sequences, collocations, have been performed (e.g.: Nesselhauf 2003, 2004; Read 2000; Schmitt and Carter 2004; Wray 2002) and the empirical research has consequently been carried out with their implications for the classroom teaching of lexis[1]. With regard to collocations in L2 learning, the use and learning of collocations seems important, for example, for the construct of an advanced L2 user who will be able to use English at an academic setting. In order to become a fluent bilingual, the use of collocations is crucial because of their pervasiveness in the L2 language.

The past research has been implemented mainly through the analysis of the types of collocations by using either written or spoken corpora so far. For example, Nation (2001) who drew up a table for listing the most frequent two- to five-word collocations occurring in the Brown Corpus demonstrated that the number of collocations occurring fifty times or more in the million word corpus is found much more in the two adjacent items than three- to five- adjacent items. From a teaching point of view, he argued that frequent collocations deserve attention in the classroom if their frequency is equal to or higher than other high-frequency words. Kilgarrif (1997) studied the frequencies of some multiword items (two- to three-words) in the British National Corpus (BNC) and found that the most frequent items marked as collocations in the BNC are counted as many as 2000.

Moreover, research into the relationships between collocations and the problems of learners of ESL (English as a Second Language) / EFL (English as a Foreign Language) were performed (e.g. Bahns 1993; Bahns & Eldaw 1993;

---

[1] For example, Lewis (2002) indicates some tips for teaching collocations for L2 learners.

Biskup 1992; Chi, *et al.* 1994; Durrant & Schmitt 2009; Farghal & Obiedat 1995; Miyakoshi 2009; Yamashita & Jiang 2010). Among them, Nesselhauf (2005) who investigated a learner written corpus of advanced German learners of English focused on the deviations in the use of collocations including verbs. She analysed various types of deviational use in collocations and discussed the use of L2 elements, L1 influence and intralinguistic factors such as the degree of combinability restriction of a collocation. Out of more than 2,000 verb-noun collocations, a quarter was found to be wrong, a third deviant. Deviation was found not only in the verbs but also in other elements of the collocations (nouns, determiners, etc.) and in the use of collocations as wholes. She argued that there are two important factors affecting the German learners' collocations: 1) congruence between their L1 equivalents in collocations and the L2; 2) the degree of restriction of a collocation, i.e. whether a verb in a collocation is restricted to a few numbers of nouns or not. The German learners were more susceptible to deviation in the collocations which were not congruent to their L1 and less restricted in combinability.

On the other hand, the types and subjects of previous studies on L2 learners' use of collocations have been limited. Previous studies discuss two types of characteristics which define collocations, i.e. combinability and transparency, one of which has been included in the empirical study but another one, transparency, has been neglected. Also, while the studies on collocations including verbs and nouns have been the main focus of the previous studies, there were not many studies on collocations consisting of other elements, such as adjective and adverbs. The subjects who have been studied previously were mainly L2 learners with Indo-European L1 backgrounds and the studies on L2 learners with non-Indo-European backgrounds have been scarce. Moreover, the learner data previously have been mainly either translation or the use of learner corpus. The study using multiple types of tasks given to learners of different L1 backgrounds on the several groups of collocations has never been performed.

Thus, the present study classified collocations based on two characteristics, combinability and transparency, in combination. The target collocations include four types of collocations, such as [verb + noun], [delexicalised verb + noun], [adjective + noun] and [adverb + adjective], which are examined with regard to the French and Japanese learners who have different L1 backgrounds.

1.2 Aims of the Study

Based on the above-mentioned backgrounds, the research in this thesis therefore analyses collocations by means of the following classifications.

Firstly, a new framework in defining collocations is provided, based on combinability and transparency of words. Then, the present study tries to classify collocations based on the two categories, combinability and transparency of words, which have often been partially discussed (See for example, Howarth 1996) but have not actually put into practice in the experimental study of collocations.

Combinability refers to the restrictedness with different words in a collocation. For example, in a collocation "make/take a decision", "make" and "take" are both combined with "a decision", and there are two possibilities of verbs to combine with a noun, "a decision". Thus, it is not restricted in the possibility of combining with different words, which are shown as [- Restricted Combinability]. When the word in a collocation has only one possibility in combining with other words, it is referred to as [+ Restricted Combinability].

Transparency can be expressed as figurativeness of the words in a collocation. It means whether the elements of the combination and the combination itself have a literal or a non-literal meaning. [+ Transparency] means the collocation has a literal meaning, while [-Transparency] refers that the collocation has a non-literal meaning. Despite the ambiguity of deciding: 1) the differences between [+Restricted Combinability] and [-Restricted Combinability]; 2) the differences between [+Transparency] and [-Transparency], the present study

attempts to provide an original framework in classifying collocations in order to bring out new findings of L2 learners' knowledge and use of collocations.

Secondly, the present study examines the knowledge and use of English collocations of different L1 backgrounds, i.e. French and Japanese learners. Among the previous studies of L2 learners' knowledge of collocations, past research has not been carried out as to compare two groups of learners with different L1 backgrounds. The present study chooses French and Japanese L2 learners to investigate how the knowledge and use of collocations are demonstrated by learners of different L1 backgrounds. Their L1s are not only different but also belong to different typological language groups: French as one of the Indo-European background languages and Japanese as one of the non-Indo-European languages.

Based on the above, the aim of the study is to investigate the differences of L1 inferences and the combinability and transparency influence of collocations between the L2 learners from different L1 backgrounds. The research questions in the present study are:

1) How different is the French and Japanese learners' L1 influence in their responses to the tasks in the present study?
2) How different are the French and Japanese learners' responses to the tasks in the present study depending on the two combinations of combinability and semantic transparency of collocations, i.e. [+Restricted Combinability, +Transparency] and [-Restricted Combinability, -Transparency]?

1.3 Outline of the Study

As shown in Figure 1.1, the present study is composed of nine chapters.

Following this introduction in Chapter 1, Chapter 2 provides a thorough overview of how to define collocation, an area where many previous vague definitions have made the concept elusive. In order to discuss the scope of collocation in the present study, two major approaches to collocation, i.e. fre-