A Multifactorial Analysis of English Particle Movement

in Korean EFL Learners' Writings

Gyu-Hyeong Lee
Hannam University
70 Hannamro, Daedeok-gu
Daejeon 306-791, Korea
gyuhyung73@naver.com

Ha-Eung Kim
Hannam University
70 Hannamro, Daedeok-gu
Daejeon 306-791, Korea
tankkh@hanmail.net

Yong-hun Lee
Chungnam Nat'l University
99 Daehak-ro, Yuseong-gu
Daejeon 305-764, Korea
yleeuiuc@hanmail.net

Abstract

This paper investigated Particle Movement in Korean EFL learners' writings. Gries (1999, 2001, 2003) adopted a multifactorial analysis to examine Particle Movement of native speakers. Several linguistics factors were proposed in the studies, and it was demonstrated that these factors influenced the choice of the constructions. This paper also employed a multifactorial analysis to examine Particle Movement in Korean EFL learners' writings. The analysis results illustrated that the Korean EFL learners were slightly different from native speakers in that only some factors were used for the selection of constructions.

1 Introduction

Linguistic alternation is one of the interesting areas in the linguistic investigations. Particle Movement is one of the syntactic alternations. It refers to the phenomenon where a particle goes behind the direct object (DO) in the phrasal verb constructions. Let's see the following example (Gries, 1999:1).

(1)   a. John *picked up* the book.
       b. John *picked* the book *up*.

In (1a), the order is 'verb + particle + DO'. However, the particle is separated from the verb in (1b). That is, (1b) has an order of 'verb + DO + particle'.

There have been a lot of theoretical studies to investigate what linguistic factors determine the choice of the alternation, in traditional grammar and generative grammar. Nowadays, as computer technology and statistics develop, there have been a few studies to explain these syntactic phenomena with corpus data. Gries (1999, 2001, 2003) adopted a multifactorial analysis to examine the Particle Movement in the native speakers' writings. These studies proposed several linguistics factors and it was demonstrated that these factors and their interactions significantly influenced the choice of the constructions

This paper also adopted a multifactorial analysis to examine the Particle Movements in Korean EFL learners' writings. The Korean part of TOEFL11 corpus was used, and all the relevant sentences were extracted using the C7 tag information. The relevant factors were encoded to these sentences, and each factor was statistically analyzed with R. Through the analysis, it was demonstrated that Korean EFL learners employed a different strategy in Particle Movement and that only some factors were used for the selection of alternation.

This paper is organized as follows. In Section 2, previous studies are reviewed with focused on corpus-based approaches. Section 3 enumerates research methods, and Section 4 contains analyses results. Section 5 is for discussions, and Section 6 summarizes this paper.

2 Previous Studies

2.1 On Particle Movement

There have been several studies on English Particle Movement in various linguistic fields: traditional grammar (Sweet, 1892; Jespersen, 1928; Kruisinga and Erades, 1953), Chomskyan transformational-generative grammar (Fraser, 1974, 1976; Den Dikken, 1992, 1995; Rohrbacher, 1994), cognitive grammar (Yeagle, 1983), discourse-functional approaches (Chen, 1986), psycholinguistically-oriented approaches (Hawkins, 1994), and so on.

Gries (1999:33) closely investigated the claims in previous studies and summarized them as follows.

Value for construction$_0$	Variable	Value for construction$_1$
Long DO	Length of the DO in words (Length W)	
Long DO	Length of the DO in syllables (LengthS)	
Complex	Complexity of the DO (Complex)	
	NP-Type of the DO: semi-pronominal (Type)	pronominal
Indefinite	Determiner of the DO (Det)	definite
No	Previous mention of the DO (Lm)	yes
Low	◀——Times of preceding mention of the DO (Topm)——▶	high
High	◀——Distance to last mention of the DO (Dtlm/ActPC)——▶	low
High	◀——News Value of the DO——▶	low
Yes	(Contrastive) Stress of the DO	
Yes	Subsequent mention of the DO (Nм)	no
High	◀——Times of subsequent mention of the DO (Tosm)——▶	low
How	◀—Distance to next mention of the DO (Dtnm/ClusSC)——▶	high
	Overall frequency of the DO (Oм)	
	following directional adverbial (PP)	yes
Yes	Prep of the following PP is identical to the particle (Part = Prep)	
	Register	
Idiomatic	◀——Meaning of the VP (Idiomaticity)——▶	literal
Low	◀——Cognitive Entrenchment of the DO——▶	high
Inanimate	Animacy of the DO (Animacy)	animate
Abstract	Concreteness of the DO (Concreteness)	concrete

Table 1. Variables That Govern the Alternation

Here, *construction$_0$* refers to the sentences with the order of 'verb + particle + DO' as in (1a), while *construction$_1$* refers to the sentences with the order of 'verb + DO + particle' as in (1b). This table enumerated 18 different linguistic factors and demonstrated that several different types of factors, not a single factor, actually influenced the choice of the constructions.

Let's see how these factors can be related with the alternation of Particle Movement. For example, LENGTHW (the first factor in Table 1) refer to the length of DO in words. If the DO is long, native speakers tend to choose *construction$_0$* rather than *construction$_1$*. If the DO is short, the native speakers prefer *construction$_1$* to *construction$_0$*. The factor DET, the fifth factor, refers to the determiner of the DO. If the determiner of DO is indefinite (such as *a* or *an*), native speakers tend to choose *construction$_0$* rather than *construction$_1$*. If the determiner of

DO is definite (such as *the*), native speakers prefer *construction$_1$* rather than *construction$_0$*. Table 1 contains all the related factors which cover most of linguistic fields: phonology, syntax, semantics, pragmatics, and discourse analysis.

2.2 Corpus-based Studies

Although it is fact that previous studies contributed to find out linguistic factors influenced the choice of alternation, their data exclusively relied on native speakers' intuition. Gries (2001, 2003) pointed out this problem and performed an analysis based on the corpus data.

Gries (2001:36-37) pointed out three problems of these previous approaches. First, most variables were based on introspective analysis and non-authentic example sentences. This problem is due to the fact that previous studies exclusively relied on the native speakers' intuition (viz. acceptability judgments). The problems of this type of test are (i) that they do not necessarily constitute objective, reliable, and valid data, (ii) that it is questionable that an analysis based on these data can in fact produce representative results, and (iii) that it is possible to evaluate sentences produced artificially, out of context. Second, most previous analyses only performed the monofactorial analyses, where only one variable has an effect on the alternation in isolation. The problem of monofactorial analysis is that the examples do not warrant the claim that the preference for one construction over the other need not be related the relevant factor exclusively. Instead, the tendency might come from other factors. Therefore, given that factors are encoded in the determination of the constructions, it is difficult to solely rely on monofactorial analyses to describe particle movement adequately. Third, there have been only a few analyses aiming at subsuming all the variables under a common basis and there has been no analysis has aimed at predicting particle placement in natural discourse situations.

In order to solve this problem, Gries (2001, 2003) employed a multifactorial analysis, where all the factors in Table 1 were taken into consideration simultaneously. These studies used a Generalized Linear Model (GLM) and statistically analyzed how each factor played a role in the choice of construction. They also took a linear discriminant analysis (LDA) and a classification and regression tree (CART) and calculated the importance of each factor as follows (Gries, 2001:48).
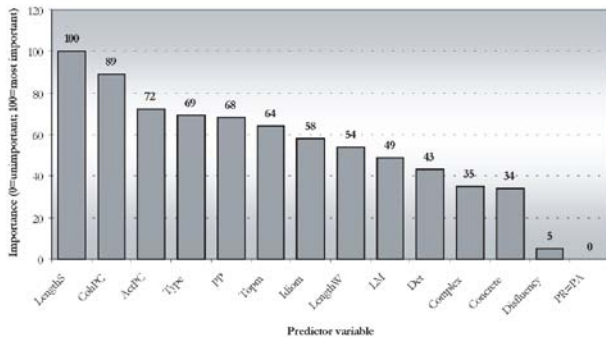
Figure 1. Importance of Predictors for CART

As this figure indicates, not all the linguistic factors play roles in the choice of alternation. In addition, some factors are more important, and others are less important.

Gries (2001) and Gries (2003) were essentially different from the previous approaches, since (i) these studies made use of corpus data (naturally occurring data) and (ii) they statistically analyzed the collected data.

3 Research Method

3.1 Questions and Hypothesis

Although there have been a lot of studies on Particle Movement in native speakers, there are few studies on the phenomena of the EFL learners. This study investigated the Particle Movement of Korean EFL learners.

Through the analysis, this paper wants to answer the following research questions.

(2) a. Do Korean EFL learners show the same or similar tendency in Particle Movement in their writings?

b. If Korean EFL learners employ different factors, which factors were employed in their choice of alternation?

c. Does the ratio of these two constructions (*construction$_0$* vs. *construction$_1$*) change as the level of proficiency goes up?

For these research questions, the following hypothesis was made.

(3) a. If Korean EFL learners show the same or similar tendency that native speakers demonstrate in their writings, two groups of people may employ similar factors or a similar set of factors in their writings that influence the choice of constructions.

b. If Korean EFL learners show a different tendency from the native speakers, two groups may employ different factors or a different combinations of factors in their writings which decide the choice of constructions.

In order to answer these questions, the following investigations were conducted.

3.2 Corpus

This study employed two types of data. The first one was the TOEFL11 corpus for the EFL learners (LDC Catalo No.: LDC2014T06), and the second one was the data in Gries (2001, 2003) for the native speakers (as reference data set). The second data were not the actual data but the analysis results in Gries (2001, 2003).

The TOEFL11 corpus was released by the English Testing Service (ETS) in 2014. The corpus consists of essays written during the TOEFL iBT® tests in 2006-2007 (Blanchard et al., 2013). It contains 1,100 essays per each of the 11 native languages, totaling 12,100 essays. All of the essays were taken from the TOEFL independent task, where test-takers were asked to write an essay in response to a brief writing topic. The essays were sampled as evenly as possible from eight different topics. The corpus also provides the score levels (Low/Medium/High) for each essay.

There are other kinds of corpora which can be used for examining the use of EFL learners. The International Corpus of Learner English (ICLE; Granger et al., 2009) is one example. Although the ICLE is a good resource to explore linguistic properties of EFL learners' use of English, there are several reasons for choosing the TOEFL11 instead of the ICLE.

First, the TOEFL11 corpus includes essays written by Korean EFL learners, while the ICLE does not.[1]

Second, each essay in the TOEFL11 corpus contains information on score levels. The score levels were calculated first by combining the individual 5-point-scale scores given by the human raters and then by collapsing this

---

[1] The ICLE corpus contains 16 components (Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Tswana, and Turkish) but the TOEFL11 includes 11 L1s (French, Italian, Spanish, German, Hindi, Japanese, Korean, Turkish, Chinese, Arabic, and Telugu).

combined score into 3 levels (Low/Medium/High). The 5-point-scale human scores were determined by the defined criteria.

Third, the number of essays that the former corpus contains is bigger than that of the latter. The ICLE includes 380 essays per L1 (=6,085/16), while the TOEFL11 has 1,100 essays per L1. Thus, the TOEFL 11 contains about three times as many essays as the ICLE.

Fourth, one of the biggest problems of the ICLE is that essay topics are not evenly distributed across the 16 L1s. The language usage is heavily driven by a given essay topic. This implies that some of the ICLE data may be conflated by the uneven distribution of essay topics across the 16 L1s. It is important to investigate the linguistic properties of EFL learners in an evenly distributed corpus.

Finally, because of the differences in the essay tasks administered and responses collected, there were differences not only in character encodings but also corpus annotations across L1s. These differences make it difficult for the findings of one L1 to be generalized to other contexts.

3.3 Procedure

The analysis in this paper proceeded as follows.

First, all the writing samples of Korean EFL learners were extracted from the TOEFL 11 corpus. A total of 328,384 word tokens were included in the extracted corpus.

Second, the writing samples were classified into three levels. Through the classification, each level had the following corpus size (word token): 95,066 (High), 202,531 (Medium), and 30,787 (Low).

Third, each text was POS tagged with the C7 CLAWS taggers.[2]

Fourth, all the sentences with particles were extracted using NLPTools (Lee, 2007).[3]

Fifth, all the relevant factors were encoded to each sentence. This paper adopted 8 factors and they are enumerated in Table 2.

Among these, the first factor was newly introduced in this analysis and the others came from Table 1.

---

[2] You can easily use Free CLAWS WWW tagger in http://ucrel.lancs.ac.uk/claws/trial.html. For details of C7 tag sets, see Jurafsky and Martin (2009).

[3] In the C7 tag sets, particles have a tag RP. The reason why NLPTools was used here is that the software had a function which could extract the whole sentences with the given tag(s) (i.e., *_RP).

Finally, all the data were statistically analyzed using R.

Variable	Explanation
LEVEL	Level of proficiency
COMPLEXITY	Complexity of Direct Object
ANIMACY	Animacy of Direct Object
DEFINITENESS	Definiteness of Direct Object
PRONOMINALITY	Pronominality of Direct Object
IDIOMACITY	Idiomacity of Direct Object
CONCRTENESS	Concreteness of Direct Object
LENGTH	Length of Direct Object in Words

Table 2. Variables Used in the Analysis

4 Analysis Results

4.1 Descriptive Statistics

After all the sentences with the particles were extracted, the sentences were classified into two groups, based on the transitive vs. intransitive use of phrasal verbs. This process was necessary since the Particle Movement occurred in the transitive use of phrasal verb constructions.

The following graph illustrates the ratio of each group (intransitive vs. transitive) of phrasal verb constructions in Korean speakers' writings.
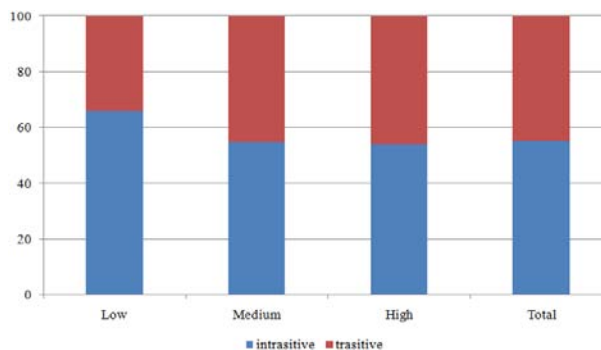

Figure 2. Intransitive vs. Transitive

Here, the lower part corresponds to the intransitive uses and the upper part to the transitive uses.

As this graph illustrated, Korean EFL learners preferred intransitive uses of phrasal verbs rather than the transitive uses. Note that nearly 50%-70% of sentences were intransitive uses of phrasal verbs. This tendency appeared in all the levels of proficiency, though the proportion of intransitive uses of phrasal verbs decreased as the level of proficiency went up. However, the proportion of the Medium level was indistinguishable from that of the High level.

Then, among the sentences with phrasal verbs, all the constructions which had transitive uses

were extracted, and the ratios of two constructions were calculated. Figure 3 illustrated the analysis results.
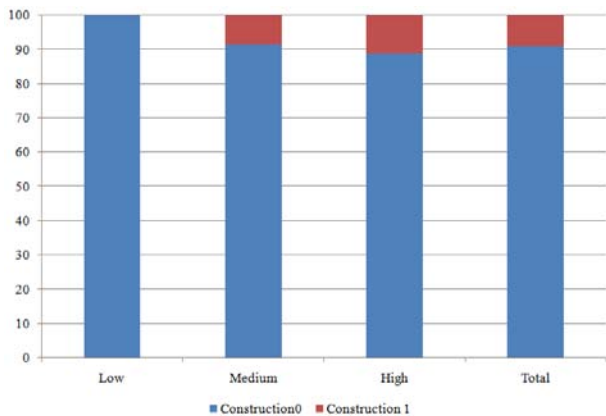

Figure 3. *construction$_0$* vs. *construction$_1$*

As this graph shows, the ratio of two constructions differ as the level goes up. The Korean EFL learners in the Low level used only *construction$_0$*. In the Medium level, the proportion of *construction$_0$* decreased and the Korean EFL learners began to use *construction$_1$*. In the High level, the proportion of *construction$_0$* decreased more, the proportion of *construction$_1$* increased more. Overall, the Korean EFL learners preferred *construction$_0$* to *construction$_1$*.

4.2   Inferential Statistics

From the data described in Figure 2, the sentences with transitive uses of phrasal verbs were extracted, since those sentences could be classified into one of the two constructions (either *construction$_0$* or *construction$_1$*). Then, a GLM was applied to the data, as in Gries (2001, 2003).

This model was chosen through the following steps. First, since we had 8 factors, a (Multiple) Linear Regression analysis is adopted (a multi-factorial analysis; Gries, 2003). Second, since the dependent variable CONSTRUCTION was binomial, a Generalized Linear Regression Model had to be used with logistic regression.

The initial model was constructed as follows.

(4)   Initial Model (Unsaturated)
CONSTRUCTION~LEVEL+COMPLEXITY+ANIMACY+DEFINITENESS+PRONOMINALITY+IDIOMACITY+CONCRETENESS+LENGTH

This is the initial model, where no interaction was included.

Then, a model selection process was performed. According to Gries (2013), there are two types of model selection parameters. One is based on the direction of the analysis and the other is the criterion determining whether or not a predictor gets to be in the model. On the direction of the analysis, most analyses have adopted a backward selection, and this paper also took this method. There are two types of approaches to the selection of relevant models: significance-based approaches and criterion-based approaches. This paper took a significance-based approach. That is, the analysis would start from the maximally saturated model, and continued to remove predictors (backward) until the analysis reached the statistically significant differences in the *p*-value (significant-based).

Since this paper adopted a backward selection, the first thing is to make a saturated model. The following model is a saturated model.

(5)   Saturated Model
CONSTRUCTION~LEVEL*COMPLEXITY*ANIMACY*DEFINITENESS*PRONOMINALITY*IDIOMACITY*CONCRETENESS*LENGTH

Note that all the interactions were included in this model.

Now that a saturated model was obtained, the statistical analysis started from the model. A new model was made by deleting one interaction or one factor from the saturated model. Then, it was checked whether this new model is significantly different from the previous model. If *p*<.05, it means that two models were significantly different and that the deleted factor or interaction MUST NOT be deleted from the model. If .05<*p*, it means that two models were not significantly different and that the deleted factor or interaction can be deleted safely without distorting the explanatory power of the model. The selection procedures were continued until no redundant factor or interaction remained in the model. Through this process, the final model was obtained.

In the final model, there were lots of interactions among the factors. Since it is impossible and unreasonable to examine all the factors and their interactions, this paper examines only the effects of major factors. The following table contains the statistical values for each factor.

	Estimate	*sd*	*z*	*p*
(Intercept)	0.702	0.611	1.150	.250
LEVEL1	0.086	0.149	0.576	.565
LEVEL2	-0.213	0.228	-0.936	.349
COMPLEXITY1	1.757	0.520	3.376	<.001
ANIMACY1	-1.090	0.114	-9.578	<.001
DEFINITENESS1	0.194	0.102	1.896	.058
PRONOMINALITY1	-0.542	0.639	-0.849	.396
IDIOMACITY1	0.132	0.087	1.508	.132
CONCRTENESS1	1.127	0.439	2.265	.010
LENGTH	0.268	0.088	3.062	.002

Table 3. Analysis Results

As this model shows, only 4 factors (COMPLEXITY, ANIMACY, CONCRETENESS, and LENGTH) among the 8 factors were statistically significant.

An interesting fact is that the factor LEVEL was not statistically significant. Though there were some differences among the level of proficiency (Figure 3), this factor LEVEL was not statistically significant as its *p*-vale indicates (*p*=.565 and *p*=.349).

4.3 Analysis with Effect Plots

Since the final model was obtained, it is possible to statistically analyze each factor and interactions with effect plots. Among the 8 factors included in the statistical analysis, only 4 main factors were statistically significant. In this section, only those 4 factors were closely examined.

The first factor is COMPLEXITY, which indicates whether the form of DO is simple or complex. Figure 4 is the effect plot for this factor.
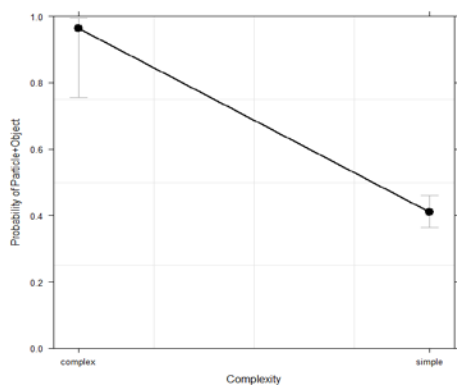
Figure 4. Effect Plot for COMPLEXITY

Compared this result with that of Gries (2003:194).
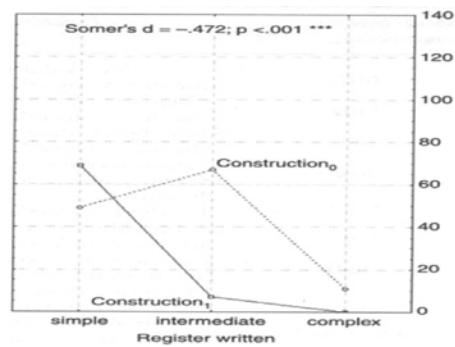
Figure 5. COMPLEX in Gries (2003)

COMPLEXITY in this paper corresponds to COMPLEX in Gries (2003).

As Figure 5 demonstrates, the *construction₀* had higher frequencies than the *construction₁* when the DO was complex. However, the *construction₁* had higher frequencies than the *construction₀* when the DO was simple. This tendency was also observed in Figure 4. When the DO was complex, the proportion of *construction₀* was greater than the value of *construction₁*, and its value was greater than 0.5. On the other hand, when the DO was simple, the proportion of *construction₁* was much greater than the value of *construction₀*, and its value was less than 0.5. Accordingly, these two graphs demonstrate that native speakers and Korean EFL learners show a similar tendency.

Since two graphs (Figure 4 and Figure 5) are different, it may be unreasonable to compare the values of two graphs. However, since the goal of comparison is to check whether the tendencies that the Korean EFL learners exhibit (not the exact values) are similar to those of native speakers, it is possible to use the analysis results in Gries (2003) in the comparison

The second factor to be mentioned is ANIMACY, which indicates whether DO was an animate or an inanimate entity. Figure 6 is the effect plot for this factor
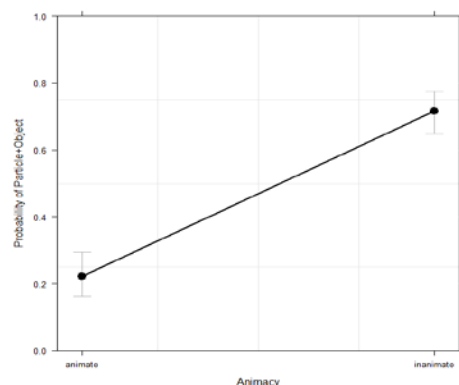
Figure 6. Effect Plot for ANIMACY

The *y* values in this graph represent the ratio of 'Particle + DO'. That is, the *y* values in this plot represent the ratio of $construction_0$. Accordingly, as the *y* value increases, the ratio of $construction_0$ increases. It means that the Korean EFL learners preferred to use $construction_0$, rather than the ratio of $construction_1$. It also implies that as the *y* value increases, the ratio of $construction_1$ decreases.
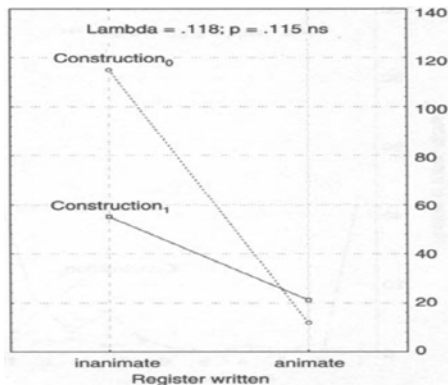
Now, let's compare this result with that of Gries (2003:197).

Figure 7. ANIMACY in Gries (2003)

ANIMACY in this paper corresponds to ANIMACY in Gries (2003). The *y* values in this graph refer to the frequencies of each construction ($construction_0$ and $construction_1$) when DO has the corresponding value for the given factor.

As Figure 7 demonstrates, though inanimate DOs were prevailed in both constructions, the $construction_1$ has higher frequencies than the $construction_0$ when the DO referred to an animate entity. However, the $construction_0$ has higher frequencies than the $construction_1$ when the DO had an inanimate entity. This tendency was also observed in Figure 6. When DO was inanimate, the proportion of 'Particle + DO' ($construction_0$) was greater than the value of 'DO + Particle' ($construction_1$), and its value was greater than 0.5. On the other hand, when DO was an animate entity, the proportion of 'DO + Particle' ($construction_1$) was greater than the value of 'Particle + DO' ($construction_0$), and its value was less than 0.5. Accordingly, these two graphs demonstrate the tendency that both native speakers and Korean EFL learners preferred to use $construction_0$ as DO took an inanimate entity.

The third factor is CONCRETENESS, which indicates whether DO refers to an abstract entity or a concrete entity. Figure 8 is the effect plot for this factor.
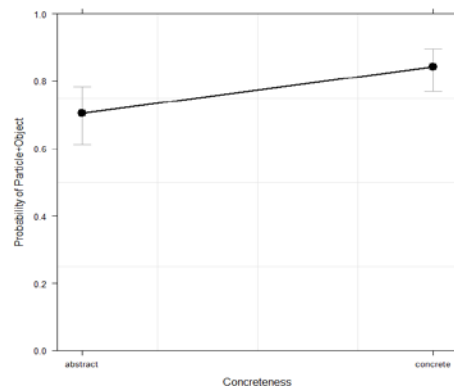
Figure 8. Effect Plot for CONCRETENESS

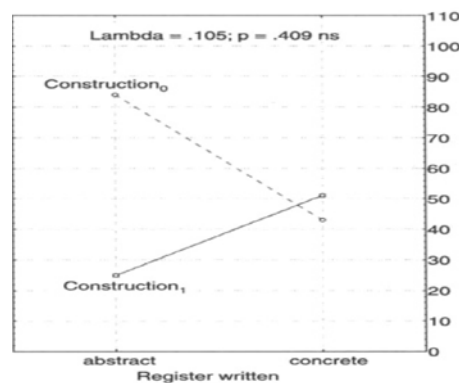Compared this result with that of Gries (2003:197).

Figure 9. CONCRETE in Gries (2003)

CONCRETENESS in this paper corresponds to CONCRETE in Gries (2003).

As Figure 9 demonstrates, the $construction_1$ has higher frequencies than the $construction_0$ when the DO had a concrete entity. However, the $construction_0$ has higher frequencies than the $construction_1$ when the DO had an abstract entity. This tendency was also observed in Figure 8. When DO had an abstract entity, the proportion of $construction_0$ was greater than the value of $construction_1$, and its value was greater than 0.5. On the other hand, when DO was a concrete entity, the proportion of $construction_1$ was much greater than the value of $construction_0$, and its value was less than 0.5. Accordingly, these two graphs demonstrate the tendency that native speakers and Korean EFL learners demonstrate an identical tendency.

The last factor to be mentioned is LENGTH, the length of DO in words. Figure 10 is the effect plot for this factor. Compared this result with that of Gries (2003:194). LENGTH in this paper corresponds to LENTHW in Gries (2003).
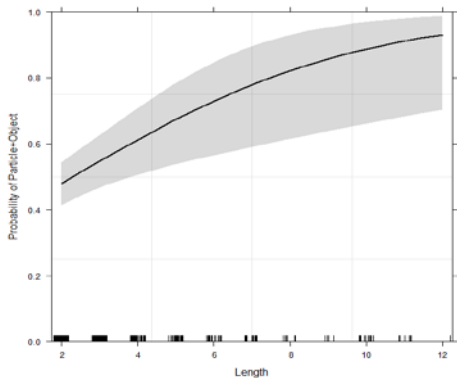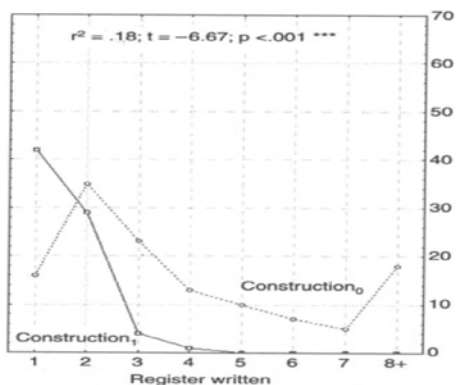
Figure 10. Effect Plot for LENGTH

Figure 11. LENGTHW in Gries (2003)

As Figure 11 demonstrates, as DO becomes longer, the *construction₀* has higher frequencies, while the *construction₁* has lower frequencies. This tendency was also observed in Figure 10. As DO becomes longer, the ratio of the order *construction₀* increases. This fact implies that the other order *construction₁* decreases. Though two graphs were slightly different, both graphs demonstrates the tendency that both native speakers and Korean EFL learners preferred to use *construction₀* as DO became longer.

5   Discussions

In Table 1, several factors were proposed which influenced the alternations of Particle Movement. Among them, 7 factors were chosen for the study in this paper: COMPLEX, ANIMACY, DET, TYPE, IDIOM, CONCRETE, and LENTHW. These factors were encoded as follows: COMPLEXITY, ANIMACY, DEFINITENESS, PRONOMINALITY, IDIOMACITY, CONCRETENESS, and LENGTH. To these 7 factors, one more factor LEVEL was added.

The comparison of Figure 1 and Table 3 demonstrated that the uses of alternation of Particle Movement in Korean EFL learners were different from those of native speakers. Among the 8 factors which influenced alternation of

Particle Movement in native speakers, only 4 factors were statistically significant in Korean EFL learners writings.

Therefore, the answer to the question (2a) will be 'No', and the answer to the question (2b) will be COMPLEXITY, ANIMACY, CONCRETENESS, and LENGTH (4 factors). As for the question (2c), there were some differences in the ratio of these two constructions as the level of proficiency goes up. However, the differences were not statistically significant.

Since Figure 1 and Table 3 demonstrated that Korean EFL learners showed different tendency in the Particle Movement in their writings compared with native speakers, the hypothesis in (3a) cannot be maintained. Instead, the hypothesis in (3b) can be supported, since a different set of factors had influenced Particle Movement in the Korean EFL learners' writings.

6   Conclusion

This paper adopted a multifactorial analysis as in Gries (2001, 2003) to examine Particle Movement in Korean EFL learners' writings. The Korean part of TOEFL11 corpus was used, and all the relevant sentences were extracted using the tag information. The eight relevant factors were encoded to these sentences, and each factor and their interactions were statistically analyzed with R.

Through the analysis, it was demonstrated that Korean EFL learners employed a different strategy in the Particle Movement and that only some factors were used for the selection of constructions. Unlike native speakers, 4 linguistic factors were statistically significant in Korean EFL learners' writing samples (ANIMACY, PRONOMINALITY, CONCRETENESS, and LENGTH). It was also observed that there were some differences in the ratio of these two constructions (*construction₀* vs. *construction₁*) as the level of proficiency went up. However, the differences were not statistically significant.

However, we do NOT say that these differences between the native speakers and the Korean EFL learners come from only the L1 transfer effects. Another kind of complicated statistical analysis (such as another regression analysis with the native data and/or the analysis in Gries and Deshors (2015)) is necessary to examine if the L1 (here, Korean) really influenced these factors and how much the L1 transfer effects are involved in these factors.

References

Daniel Blanchard, Joel Tetreault, Derrick Higgins, Aoife Cahill, and Martin Chodorow. 2013. TOEFL11: A corpus of non-native English. ETS RR–13-24. Prinston, NJ: Educational Testing Service.

Ping Chen. 1986. Discourse and Particle Movement in English. Studies in Language 10:79-95.

Marcel Den Dikken. 1995. Particles: On the Syntax of Verb-Particle, Triadic, and Causative Constructions. Oxford: Oxford University Press.

Marcel Den Kikken. 1992. Particles. Holland Institute of Linguistics Dissertations. The Hague: Holland Academic Graphics.

Bruce Fraser. 1974. The Phrasal Verb in English, by Dwight Bolinger. Language, 50:568-575.

Bruce Fraser. 1976. The Verb-Particle Combination in English. New York: Academic Press.

Sylviane Granger, Estelle Dagneaux, and Fanny Meynier. 2009. The international corpus of learner English: Handbook and CD-ROM (version 2). Louvain-la-Neuve, Belgium: Presses Universitaires de Lowvain.

Stephan Th. Gries. 1999. Particle movement: A Cognitive and Functional Approach, Cognitive Linguistics, 10(2):105-145.

Stephan Th. Gries. 2001. A Multifactorial Analysis of Syntactic Variation: Particle Movement Revisited. Journal of Quantitative Linguistics, 8(1):33-50.

Stephan Th. Gries. 2003. Multifactorial Analysis in Corpus Linguistics: A Study of Particle Movement. London: Continumm.

Stephan Th. Gries. 2013. Statistics for Linguistics with R: A Practical Introduction. Berlin: Guyter.

Stephan Th. Gries and Sandra Deshors. 2015. EFL and/vs. ESL? A Multi-level Regression Modeling Perspective on Bridging the Paradigm Gap. International Journal of Learner Corpus Research 1(1): 130–159.

John Hawkins. 1994. A Performance Theory of Order and Constituency. Cambridge: Cambridge University Press.

Otto Jespersen. 1928. A Modern English Grammar on Historical Principles. London: George Allen and Unwin Ltd.

Jurafsky, Daniel and James Martin. 2009. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Upper Saddle Hill, NJ: Prentice Hall.

Etsko Kruisinga and Patrick Erades. 1953. An English Grammar. Vol. I. Groningen: P. Noordhoff.

Yong-hun Lee. 2007. Corpus Analysis Using NLPTools and Their Applications: Applications to Linguistic Research, English Education, and Textbook Evaluation. Seoul: Cambridge University Press.

Bernhard Rohrbacher. 1994. English Main Verbs Move Never. The Penn Review of Linguistics, 18:145-159.

Henry Sweet. 1892. A New English Grammar. Oxford: Clarendon Press.

Rosemary Yeagle. 1983. The Syntax and Semantics of English Verb-Particle Constructions with off: A Space Grammar Analysis. Unpublished M.A. Thesis, Southern Illinois University at Carbondale.