『영미연구』
제51집 (2021): 223-250
http://doi.org/10.25093/ibas.2021.51.223

The Effects of Verb Bias in the Written Production of English Sentences by L2 Learners*

Ji Hyon Kim
Hankuk University of Foreign Studies

[Abstract]

A written production task was used to examine the contributions of frequency-based verb bias information to L2 learners' production of English sentences. Korean L2 learners of English were instructed to produce a complete sentence for each of forty-three English verbs. The verbs belonged to three categories based on verb bias: direct object (DO) bias, sentential complement (SC) bias, and equibias (EQ-bias). The proportion of DO sentence completions was highest for the DO-bias verbs and lowest for the SC-bias verbs, replicating and extending results found in previous studies with monolingual English speakers. The results also showed a significant positive correlation between the proportion of DO/SC sentence structures for each verb in this study and the corresponding proportion of DO/SC sentence completions reported in a previous study with native English speakers (Garnsey et al., 1997). These results imply that the L2 learners in this study were sensitive to the continuous nature of

* This work was supported by Hankuk University of Foreign Studies Research Fund 2021.

verb bias and are able to employ this information in their written production. Methodological issues regarding the treatment of verb bias information as a categorical variable in psycholinguistic studies are also discussed.

Key Words: verb bias, frequency-based information, norming, written production task, L2 learners

1. Introduction

Language processing requires the rapid integration of each incoming word into the existing phrase structure based on our knowledge of the syntactic rules and probabilistic frequencies of different structures in the language. During this process, there may exist various points in the sentence at which more than one possible relationship is possible among the words, leading to temporary ambiguity. For example, as shown in (1), a sentence beginning with *The girl feared the lion ...* can be completed in two different ways depending on the grammatical role of *the lion*.

(1) The girl feared the lion ...
  a. ... at the zoo.
  b. ... would escape from its cage.

In (1a), the NP *the lion* is the direct object of the main verb *feared*. In (1b), *the lion* does not play the role of direct object, but is the subject of the embedded sentential complement. If a reader analyzes *the lion* as the direct object, predicting a continuation of the sentence similar to (1a), but finds that the sentence continues as

(1b), the reader will then have to backtrack and reanalyze the sentence. This type of reanalysis of a temporary ambiguity is called a "garden-path" (Frazier, 1978; Frazier & Fodor 1978; Frazier & Clifton, 1996). The ambiguity found in sentences such as (1) is referred to in the psycholinguistic literature as the "direct object (DO) vs. sentential complement (SC) ambiguity," and poses difficulties for the reader when incorrect predictions are made, as reanalysis of a sentence will result in increased time and processing costs.

Fortunately, we are not "garden-pathed" every time we encounter a DO/SC ambiguity because verbs differ in the relative frequencies of the linguistic structures that they occur in. For example, some verbs, such as *accept* occur with higher frequency in a DO structure, while other verbs such as *admit* are found more frequently in a SC structure. This type of information is called verb bias, and is employed in language processing to predict how incoming words should be integrated into the existing structure and decrease the probability of a reanalysis. Research has shown that verb bias information not only affects processing difficulty (reflected as higher reading times) in sentence processing, but sentence production as well (Gahl & Garnsey, 2004).

Although knowledge of verb bias plays a substantial role in the processing and production of language, most of the existing studies on the role of verb bias has been centered on the monolingual English population. Research on L2 learners has mostly been focused on whether L2 learners have successfully acquired knowledge of verb transitivity. A multitude of studies with L2 learners have examined whether they are able to distinguish transitive and intransitive verbs and use them in the appropriate syntactic contexts (Balcom, 1997; Montrul, 2001; Ju, 2000). In contrast, few studies in the L2 literature have explored the issue of whether L2 learners are sensitive to differences in verb bias frequency in English verbs.

This study investigates whether L2 learners of English are affected by the probabilistic frequencies with which English verbs occur in different syntactic structures in their written production and compares L2 learners' use of verb bias frequency to that of English monolinguals. By comparing the L2 learner data to previous norming studies and corpus data from monolingual English speakers (Biber, Conrad, & Reppen, 1998; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Roland & Jurafksy, 2002), the results of this study may serve as a foundation for future psycholinguistic studies on L2 processing and disambiguation of sentences with temporary ambiguity.

2. Literature Review

English verbs are categorized as transitive or intransitive based on whether they may be used with a direct object or not. In psycholinguistics, transitive verbs may be divided into subtypes based on what type of complement they most frequently occur with. Depending on the type of linguistic structure that verbs occur in with greater frequency, verbs are classified into DO-bias verbs vs. SC-bias verbs. Experimental studies have shown that readers actively employ verb bias information during sentence processing. Consequently, sentence structures that are consistent with a verb's bias are easier to process, whereas sentences that do not match verb bias information will cause more difficulty for the reader (Gahl, 2002; Garnsey et al., 1997; Hare, McRae, & Elman, 2003; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell, Tanenhaus, & Kello, 1993). The successful acquisition of verb bias information enables the reader to predict incoming linguistic input and integrate it with the existing syntactic structure rapidly and efficiently. The match between verb

bias and corresponding sentence structure affects the processing difficulty of a sentence to such a large degree that it may override other factors, such as sentence complexity, also reported to affect processing difficulty (Gahl, Menn, Ramsberger, Jurafsky, Elder, Rewega, & Holland, 2003).

Although most existing studies in the field of L2 acquisition have been focused on L2 learners' successful acquisition and use of verb transitivity (Brooks, Tomasello, Dodson, & Lewis, 1999; Montrul, 2001; Oshita, 2000), a few studies have recently begun to investigate the role of verb bias in L2 learners. In a self-paced reading study, Lee, Lu, and Garnsey (2013) tested whether L2 learners were sensitive to verb bias information during sentence processing. The results showed that only the participants with higher L2 proficiency were able to use complementizer and verb bias cues to predict upcoming material during sentence processing, suggesting that L2 proficiency may be a factor that interacts with L2 learners' ability to make use of verb bias information. Kaan (2014) extended these findings by presenting evidence that additional individual factors in L2 learners including differences in frequency biases, the accuracy and consistency of lexical representations, and task effects may also affect how verb bias information is used by L2 learners. Individual factors such as L2 proficiency or differences in frequency biases may affect the use of verb bias information to a greater degree than verb transitivity information because deviating uses of verb transitivity will result in ungrammaticality, leading to correction or feedback, while differences in use of verb bias will not.

The few L2 studies on verb bias that do exist make use of verb bias frequency data that is collected from L1 corpora or L1 norming studies. However, there are several issues that must be addressed before using verb bias frequency information obtained from existing L1 resources on native English speakers in the construction of experimental material in L2 processing studies.

In English monolingual research, two types of resources have generally been used to provide information on verb bias frequencies as a base for constructing experimental stimuli in processing studies. The first type of resource is corpus data, in which verb bias frequencies are extracted from naturally-occurring speech (Marcus, Santorini, & Marcinkiewicz, 1993; Macleoad, Grishman & Meyers, 1997; Lapata, Keller, & Schulte im Walde, 2001). The second type of resource provides verb bias frequencies based on experimental norming studies. Norming studies usually take the form of written sentence completion tasks, in which monolingual English speakers are provided with sentence-initial fragments, usually composed of an 'NP + Verb' sequence. Based on the proportion of sentence completions with a DO or SC structure, the verb is categorized as either a DO-bias or SC-bias verb.

One potential problem with the classification of verbs into DO-bias vs. SC-bias verbs lies in the classification criteria that is used. Some studies use the absolute classification method by which a verb is classified as DO-bias if it takes a DO complement more than 50% of the time (Lapata et al., 2001; Merlo, 1994). Other studies use the relative classification method and classify a verb as DO-bias if it is used in a DO structure at least twice as often as it is used in a SC structure (Garnsey et al., 1997; Pickering & van Gompel, 1996; Pickering, Traxler, & Crocker, 2000; Trueswell et al., 1993). Although the results obtained through the absolute and relative classification criteria generally converge, on some occasions, the two methods will result in opposite verb bias classifications, especially in the case of SC-bias verbs, as SC completions are generally less frequent across verbs compared to DO completions.

A second problem with the existing monolingual resources on verb bias frequencies lies in the categorical, split-group analysis. Most previous datasets categorize verbs into two or three subgroups: DO-bias, (Equibias,) and SC-bias verbs.

However, whether using naturally-occurring corpus data or the results of experimental norming studies as the base for verb bias classification, verb bias is a continuous variable that is calculated based on the relative frequencies of DO and SC usage, and not a categorical variable. Conducting a categorical analysis of verb bias when verb bias frequencies are of a continuous nature in reality may obscure graded differences among verbs in the same category. For example, the verb *accept* was used with a DO complement for 98% of the sentence completions in a norming task by Garnsey et al., (1997), while the verb *protest* was used with a DO complement in 58% of the responses. In spite of this difference in the proportion of DO usage, both verbs were classified as DO-bias verbs. In fact, the proportions of DO completions for the verbs classified as DO-bias in Garnsey et al. (1997) range from 0.45 to 0.98, while the proportions of SC completions for the SC-bias verbs span an even wider range, from 0.24 to 0.90. When verbs that are distributed along the verb bias continuum are categorically divided into two or three groups and each group is treated as having homogeneous qualities, the resulting data will not accurately reflect how verb bias information is accessed and employed by readers during sentence comprehension.

A third problem with adopting verb bias frequencies from existing monolingual studies and using them as a base for the construction of experimental stimuli for L2 studies concerns ecological validity. Exposure-based accounts claim that the probabilistic frequencies with which we are exposed to specific syntactic structures in the daily linguistic input significantly affect the parsing strategies that we employ during sentence processing (Dussias & Sagarra, 2007; Frenck-Mestre & Pynte, 1997). Therefore, corpus data and experimental norming data obtained from monolingual speakers in an L1 environment may not be ecologically valid for L2 learners. The different linguistic environment that L2 learners are exposed to may also entail different frequencies of exposure to the relative syntactic constructions under interest.

For example, a verb that is considered SC-bias by monolingual English speakers may be considered DO-bias for L2 learners if that particular verb has been encountered more frequently within a DO structure in their daily L2 input.

The goal of the present study is to address the issues discussed above concerning the existing resources on verb bias frequencies in the monolingual literature and investigate whether these resources are viable for L2 learners. In this study, Korean L2 learners of English participated in an elicitation task designed to investigate how verb bias information is used during written production. The verbs used as experimental stimuli were selected so that the absolute and relative classification criteria did not yield conflicting results as to verb bias categorization. Furthermore, analysis of the written production data was carried out through a series of correlation analyses, so that verb bias frequencies were treated as a continuous variable, and not a categorical variable. The results of the present study are expected to contribute to the validity of future research on verb bias frequencies and provide a database that may be used when designing experimental stimuli for L2 processing studies. The research questions are presented below.

(2a) Are L2 learners of English able to use verb bias information as a continuous variable in their written production? Does the L2 production data match the verb bias frequencies reported for English monolinguals in previous L1 research?

(2b) Is L2 proficiency a factor that affects the use of verb bias frequencies found in the L2 production data?

3. Research Method

3.1 Participants

Thirty-seven undergraduate students at a Korean university were recruited for the present study. Before the main experiment was conducted, the participants took an English proficiency test designed to assess proficiency in English reading, writing, grammar, and vocabulary. After the proficiency test, the participants completed a language background survey which included a self-rating of their general English proficiency in addition to English proficiency skills in reading, writing, listening and speaking. The survey also collected information regarding the participants' native language, any languages spoken other than English, and the length of time spent in foreign countries. Three participants who reported having a native language that was not Korean (two Chinese, one English) and one participant whose native language was Korean but had spent three years in an English-speaking country were excluded from the main experiment. These participants were not included in the main experiment as the different linguistic environments in which the participants had been exposed to English may have affected the acquisition of verb bias information. This left a total of thirty-three participants who took part in the written production task. A summary of the background information of these participants is presented in Table 1.

Table 1. Summary of L2 Participants' Information

	Age	L2 proficiency[1)]	Self-rated proficiency[2)]				
			R	W	S	L	O
Mean	23.86	23.03	3.90	2.96	2.57	3.74	3.56
SD	2.74	3.81	0.61	0.92	0.74	0.89	0.58
Range	21-27	15-28	3-5	2-5	2-5	2-5	2-5

3.2 Experimental Materials

The English verbs used as experimental stimuli in the written production task were selected from the verb bias norming study by Garnsey et al. (1997). In this study, Garnsey et al. (1997) conducted a sentence completion task with English monolingual speakers and used a relative classification criteria to subcategorize verbs into DO-bias, Equibias, and SC-bias verbs. Among the 48 verbs originally used in Garnsey et al. (1997), five verbs that resulted in contrasting classifications when using the absolute classification method were excluded from the present study. This resulted in a total of 43 English verbs (14 DO-bias, 15 Equibias, and 14 SC-bias). The 43 verbs represented the continuum of verb bias frequencies, with the proportion of DO completions obtained for each verb ranging from 0.01 to 0.98 and proportion of SC completions ranging from 0.01 to 0.89. The verbs were matched in length and frequency to control for factors other than verb bias information that may affect the results. The properties of the verbs that were used in the written production task are presented in Table 2.

Table 2. Properties of Experimental Verbs Used

	DO completions	SC completions	Frequency	Length
DO-bias verbs	77%	13%	146	8.2
Equibias verbs	36%	37%	178	7.6
SC-bias verbs	11%	59%	128	7.7

The 43 English verbs were pseudo-randomized in the written production task so that no more than two verbs from the same verb bias category were presented consecutively. Instead of presenting sentence-initial fragments composed of a proper name followed by a verb (e.g., *Debbie remembered* _____) as in Garnsey et al. (1997), the written production task in the present study simply presented each verb and required the participants to write a complete sentence with the given word. This methodology was chosen to allow for a broader range of sentence structures to naturally occur in the resulting data. For example, the passive structure would not be a possible option if participants are required to provide sentence completions for "Subject + Verb" beginnings, in which the verb is in the active past-tense form. In order to minimize the number of sentences in which the given verb was used as a different word category, the instructions specified that the presented words were all verbs.

3.3 Experimental Procedure

After the participants had taken the English proficiency test and completed the language background information survey, the written production task was conducted in a quiet classroom. English instructions were given on the test form and also explained orally if participants asked for clarification. The participants were instructed

to provide a complete, grammatical sentence for each of the 43 English verbs. The entire experimental procedure took approximately 45 minutes.

3.4 Data Coding and Analysis

The criteria used to code the written production data was based on the coding framework provided by Gahl, Jurafsky, and Roland (2004). Sentence structures in which the verb was followed by a direct object complement were coded as *DO* and verbs followed by sentential complements were coded as *SC*. Sentence completions in which a 'particle + NP object' followed the verb were coded as *PT,* and sentence completions with an infinitival form of a verb as the complement were coded as *INF*. Verbs followed by an adjectival or adverbial complement were coded as *AD*. Passive structures were coded as *PAS*, and sentences in which the verb was used ungrammatically were coded as *UNG*. Sentences that included grammatical errors that were not related to verb usage (such as incorrect use of determiners) were not coded as UNG. Finally, sentences in which the participant used the presented word as some other grammatical category instead of verb were coded as *Other*. A summary of the coding criteria including sample sentences for each category taken from the written production data is presented in Table 3.

Table 3. Coding Criteria

Sentence Type	Sample Sentence
DO (NP direct object)	*The man established **an organization**.*
SC (Sentential complement)	*She guaranteed **that you were right**.*
PT (Particle + NP)	*I believed **in you.***
INF (Infinitival verb)	*He decided **to go** there.*
AD (Adjectival or adverbial)	*I felt **cold** yesterday.*
PAS (Passive)	*The theory was proved incorrect.*
UNG (Ungrammatical)	**I feared of goldfishes.*
Other	*My mother is always worried about me.*

After the data were coded, a series of correlation analyses were conducted to examine the relationship among different variables. First, in order to examine whether the L2 learners' use of verb bias in their written production patterned with that of English monolinguals, a correlation analysis was conducted between the proportion of DO/SC sentences for the L2 learners and those of the monolingual English speakers in Garnsey et al. (1997). Next, the effects of L2 language proficiency were investigated by entering each participant's English proficiency score and their proportion of DO/SC sentence completions into a correlation analysis. The results of these analyses are reported in the following section.

4. Results

A total of 1,419 sentences were collected as data from participants in the written production task. A summary of the coded production data by verb type is presented in Table 4, followed by a descriptive analysis.

Table 4. Number of Sentence Structures by Verb Type
(percentage of total occurrences)

	DO-bias	Equibias	SC-bias	Total
DO (NP direct object)	293	207	113	613 (43.20%)
SC (Sentential complement)	71	142	193	406 (28.61%)
PT (Particle + NP)	31	14	67	112 (7.89%)
INF (Infinitival verb)	6	10	44	60 (4.23%)
AD (Adjectival or adverbial)	2	17	2	21 (1.48%)
PAS (Passive)	83	56	26	165 (11.63%)
UNG (Ungrammatical)	26	45	48	119 (8.39%)
Other	4	11	11	26 (1.83%)

Sentence structures with a DO complement and sentential complement (SC) occurred most frequently across the collected written production data and accounted for 71.81 % of the data (1019 occurrences). In general, participants produced more DO sentence structures compared to SC sentence structures (613 vs. 406 occurrences), but the relative proportion of DO vs. SC sentences varied by verb type. DO sentence structures were used most frequently for verbs categorized as DO-bias verbs in Garnsey et al. (1997) and least frequently for SC-bias verbs. In contrast, SC sentence structures were used most frequently for SC-bias verbs and least for DO-bias verbs. Passive sentence structures showed a pattern similar to that of DO sentence structures, with the highest proportion of passive sentences found with DO-bias verbs and least with SC-bias verbs.

Sentence structures in which the verb was followed by a 'particle + NP' sequence occurred most frequently for SC-bias verbs. This type of complement was found to be focused on a few particular verbs, rather than being evenly distributed throughout the experimental stimuli. Three verbs belonging to the SC-bias category (*argued about/with, worried about, figured out*) and one Equibias verb (*protested about*)

constituted 72 of the 112 occurrences (64.29%). Infinitival verb complements made up 4.34% of the entire dataset and occurred most frequently for the verb *decide* (29 out of 60 occurrences). 21 sentences used verbs with adjective/adverbial complements, and 17 out of these 21 sentences (80.95%) occurred with a single verb, *feel*. 26 sentences in which the verb was used as a category other than the main verb were found. These sentences mostly used the past participle form of the verb as a modifier or complement, and were relatively equally distributed among the verbal stimuli. Ungrammatical sentences due to incorrect use of the verb accounted for 8.39% of the data, and were found in greater proportion for SC-bias and Equibias verbs compared to DO-bias verbs.

Next, in order to answer the main research question (2a), and investigate whether the L2 participants' usage of verb bias information matched the existing resources with monolingual English speakers, a correlation analysis was conducted with the proportion of DO/SC sentence structures for each verb obtained from the present study and the DO/SC structures for the same verbs as reported in Garnsey et al. (1997). The results yielded a significant positive correlation between DO sentence structures in the L1 and L2 data (r=.58, p<.0005). A significant positive correlation was also found between the proportion of SC sentence structures in the L1 and L2 data (r=.51, p<.0005). In the monolingual (L1) norming data, correlations between the relative frequencies of DO and SC sentence structures showed a significant negative correlation (r=-.50, p=001). However, the proportions of DO and SC sentence structures in the L2 data for the present study were not significantly correlated (r=-.08, p=.63).

In order to investigate the effect of English proficiency on written sentence production, each participant's proficiency score was entered into a correlation analysis with the proportion of DO structures, proportion of SC structures, and number of

ungrammatical sentences for each participant. The results of this analysis showed a significant positive correlation with English proficiency and proportion of SC sentence structures that were used (r=.37, p=.04), indicating that participants with higher proficiency tended to use SC structures more often. English proficiency scores, however, were not significantly correlated with the proportion of DO sentence structures or ungrammatical sentences found in the data (all ps > .20).

5. Discussion

The present study was designed to investigate whether Korean L2 learners' were able to use verb bias information in their written production, and whether the relative frequencies of sentence structures found for the L2 learners matched those reported in the monolingual English speakers. In contrast to previous studies which have generally employed a categorical grouping method of verbs into DO-bias, Equibias, and SC-bias verb categories (Garnsey et al., 1997; Lapata et al., 2010; Merlo, 2003; Pickering et al., 2000; Trueswell et al., 1993), the present study focused on the continuous nature of verb bias frequencies. To this end, correlation analyses treating verb bias as a continuous variable were performed on the data in addition to a descriptive analysis with the traditional categorical split-group method. This type of experimental design has advantages in that it is able to reflect more accurately how verbs are used in naturally-occurring language by taking into account the entire continuum of frequencies of DO and SC usage of verbs.

An additional feature of the experimental design used in the present study was the presentation of single verbs instead of the sentence-initial fragments generally used in most norming studies (Garnsey et al., 1997; Pickering et al., 2000). While sentence

completion studies have advantages in that the resulting completions are mainly limited to the DO/SC transitive structure or the intransitive structure, this type of design also has limitations in that it sharply restricts the options available to the participant. Therefore, sentence completion tasks may yield production data that diverge from the frequencies of sentence structures found in naturally-occurring language. In the present study, the participants were instructed to use each given verb in a sentence and given the freedom to choose any sentence structure of their choice. As a result, although DO/SC structures comprised the greater part of the data (71.81%), a number of passive sentence structures were also observed (11.63%). The passive structure is not an option for participants with the sentence-initial fragment norming study used by Garnsey et al. (1997), suggesting that this type of experimental design more accurately reflects the diversity of sentence structures that are observed in naturally-occurring language[3)].

In order to answer the question of whether L2 learners are sensitive to the continuous nature of verb bias, both descriptive analyses and correlational analyses were conducted on the collected written production data. The descriptive analysis of the data suggests that the L2 learner participants had acquired verb bias information and were able to employ this information in their written production. The greatest number of DO sentence structures was obtained for the verbs categorized as DO-bias in Garnsey et al. (1997). SC-bias verbs showed the lowest proportion of DO sentences, with Equibias verbs showing a proportion between DO-bias and SC-bias verbs. The opposite pattern was found for SC sentence structures. The proportion of SC structures used in the written production data was largest for the SC-bias verbs and smallest proportion for DO-bias verbs.

The correlation analyses were conducted in order to examine whether the L2 learners were sensitive to the continuous nature of verb bias frequencies. Significant

positive correlations were found between the proportion of DO and SC sentence structures found in the L2 data and the DO and SC frequencies reported in the monolingual norming data (Garnsey et al., 1997, Gahl et al., 2004). These results indicate that the relative proportion of DO/SC structures used for each verb by the L2 learners in this study matched the probabilistic frequencies of verb bias reported in the monolingual L1 data. Instead of adopting a categorical grouping method and treating all verbs within the same category identically, the L2 learners' preference for a DO sentence structure increased corresponding to the actual frequencies with which the DO structure was found in naturally-occurring speech data. These results indicate that the L2 learners were indeed sensitive to the continuous nature of verb bias and employed this knowledge in their written production.

The significant correlations found between L2 learners' usage of verb bias and the probabilistic frequencies reported in the L1 literature contribute toward resolving the issues previously discussed regarding the existing data on English verb bias information. First, divergent results from using absolute and relative classification criteria to categorize verbs need not be an issue if future studies take into account the continuous nature of verb bias in experimental design and analysis of the data, instead of employing a split-group categorization method. In addition, if L2 learners are sensitive to the relative frequencies of verb bias as reflected in the monolingual L1 corpus data, the ecological validity of using existing L1 corpus or norming data as a baseline for designing the experimental stimuli in L2 studies may not be a source of concern.

However, there remains the possibility that L2 proficiency could be a factor that affects the use of verb bias information in written production, which leads to the second research question. A significant positive correlation was found between each participant's English proficiency score and the proportion of SC sentence structures

found in their written production data. These results indicate that the proportion of SC sentence structures that each participant used corresponded to their level of English proficiency, so that participants who were highly proficient in English used more SC sentence structures.

These results may be understood in terms of syntactic complexity and our general preference for simple structures. Models of sentence processing such as the Garden-path Model (Frazier, 1978; Frazier & Clifton, 1996) claim that when faced with a temporary ambiguity in a sentence, we follow the Minimal Attachment Principle, which chooses the syntactic structure with the fewest possible nodes. When comparing the DO and SC sentence structures, the SC structure is the more complex, as more nodes are required in the syntactic structure in order to instantiate a new embedded clause. Therefore, when given the option between the two structures in a neutral context, it follows that more L2 learners would choose the DO structure, which poses less of a cognitive processing load. However, when the verb bias frequencies acquired by the L2 learner for a particular verb strongly support a preference for the SC structure, the inclination toward the simple DO structure may be overridden, resulting in the positive correlation between SC bias frequencies and the actual number of SC structures found in the production data. It may be the case that when L2 proficiency is at a lower level, the L2 learner chooses the more efficient option requiring less of a cognitive processing load, i.e., the DO sentence structure, in spite of the contrasting verb bias information. This will lead to the resulting positive correlation between L2 proficiency and SC structures, with more proficient L2 learners using the SC sentence structure more often.

The number of ungrammatical sentences found for each verb type reported in the descriptive analysis supports these claims regarding L2 learners' general inclination towards the most efficient, simplest structure. The least number of ungrammatical

sentences resulting from incorrect use of the verb was found for DO-bias verbs. In contrast, approximately twice the number of ungrammatical sentences was found for Equibias and SC-bias verbs. These results suggest that when the verb bias frequencies acquired for a particular verb indicated a preference for the syntactically more complex SC structure, the L2 learners attempted to produce the required structure, but resulted in a higher proportion of grammatical errors in the process.

The correlation analysis between English proficiency scores and number of grammatical errors did not yield a significant correlation, suggesting that participants with lower levels of proficiency did not necessarily make more errors. The lack of a significant correlation may have been due to the fact that in general, the L2 participants in the present study were a relatively homogeneous group, ranging from intermediate to high levels of L2 proficiency, as all of them were undergraduate students majoring in English. As a result, the proportion of grammatical errors found in the written production data was not that high (8.39% of the entire dataset) and occurred on average in only approximately 3.6 sentences out of the 43 sentences provided by each participant. Further studies including L2 learner participants with lower levels of proficiency are expected to shed more light on this topic.

6. Conclusion

In the present study, Korean L2 learners' usage of verb bias information was investigated through a written production task. The results found suggest that the L2 learner participants were able to employ this information in their written production to reflect the continuous nature of verb bias frequencies found in naturally-occurring language data. These findings provide a significant contribution towards future studies

in L2 processing by establishing that L2 learners are able to acquire linguistic competence related to the probabilistic frequencies of verb bias to a degree that matches monolingual L1 speakers. In addition, these results may alleviate previous concerns regarding the ecological validity of using L1 corpus data or data from L1 norming studies as the baseline for constructing L2 experimental stimuli.

However, the present study also found English proficiency to be a factor that may affect the relative frequencies with which the more complex SC structure is used by L2 speakers. Future studies with L2 learners must take into consideration the fact that the L2 learner population is much more diverse than the L1 monolingual population, both in terms of L2 proficiency and the amount and nature of L2 exposure. Therefore, instead of assuming that L2 learner participants' exposure to verb bias frequencies will always match the frequencies of sentence structures previously reported in L1 databases, the experimenter should always confirm these assumptions through a norming study prior to the main experiment. Further studies employing L2 learner participants from a wider range of proficiency levels are expected to provide a clearer view on the effects of L2 proficiency on the acquisition and use of verb bias on L2 sentence processing and production.

Notes

1) Self-rated proficiency was rated on a scale of 1 to 5 (from 'not proficient at all' to 'very proficient'). O (overall proficiency), R (Reading), W (Writing) S (Speaking), L (Listening) and was used to complement the main English proficiency test.

2) The English proficiency test had a total of thirty questions and was obtained from the Department of English at the University of New Hampshire. A full copy of the test may obtained by contacting the author.

3) The proportion of passive structure in nontechnical discourse is reported to be approximately 13% (Gahl et al., 200; Givon 1979).

Works Cited

Balcom, Patricia. "Why is This Happened? Passive Morphology and Unaccusativity." *Second Language Research*, vol. 13, no.1, 1997, pp. 1-9.

Biber, Douglas, Susan Conrad, and Randi Reppen. *Corpus Linguistics: Investigating Language Structure and Use*, edited by Randi Reppen and Susan Conrad. Cambridge UP, 1998.

Brooks, Patricia, Michael Tomasello, Kelly Dodson, and Laurence Lewis. "Young Children's Overgeneralizations with Fixed Transitivity Verbs." *Child Development*, vol. 70, no. 6, 1999, pp. 1326-37.

Dussias, Paola, and Nuria Sagarra. "The Effect of Exposure on Syntactic Parsing in Spanish-English Bilinguals." *Bilingualism: Language and Cognition*, vol. 10, no.1, 2007, pp. 101-16.

Frenck-Mestre, Cheryl, and Joel Pynte. "Syntactic Ambiguity Resolution While Reading in Second and Native Languages." *The Quarterly Journal of Experimental Psychology*, vol. 30, no. 1, 1997, pp. 119-48.

Frazier, Lyn. "On Comprehending Sentences: Syntactic Parsing." Doctoral Dissertation. University of Connecticut, 1978.

Frazier, Lyn, and Charles Clifton. *Construal*. MIT P, 1996

Frazier, Lyn, and Janet Fodor. "The Sausage Machine: A New Two-Stage Parsing Model." *Cognition*, vol. 6, no. 4, 1978, pp. 291-325.

Gahl, Susanne. "The Role of Lexical Biases in Aphasic Sentence Comprehension." *Aphasiology*, vol. 16, 2002, pp. 1173-98.

Gahl, Susanne, Dan Jurafsky, and Douglas Roland. "Verb Subcategorization Frequencies: American English Corpus Data, Methodological Studies, and Cross-Corpus Comparisons." *Behavior Research Methods, Instruments, and*

Computers, vol. 36, no. 3, 2004, pp. 432-43.

Gahl, Susanne, and Susan Garnsey. "Knowledge of Grammar, Knowledge of Usage: Syntactic Probabilities Affect Pronunciation Variation." *Language*, vol. 80, 2004, pp. 748-75.

Gahl, Susanne, Lise Menn, Gail Ramsberger, Daniel Jurafsky, Elizabeth Elder, Molly Rewega, and Audrey Holland. "Syntactic Frame and Verb Bias in Aphasia: Plausibility Judgments of Undergoer-Subject Sentences." *Brain and Cognition*, vol. 53, 2003, pp. 223-28.

Garnsey, Susan, Neal Pearlmutter, Elizabeth Myers, and Melanie Lotocky. "The Contributions of Verb Bias and Plausibility to the Comprehension of Temporarily Ambiguous Sentences." *Journal of Memory and Language*, vol. 37, 1997, pp. 58-93.

Hare, Mary, Ken McRae, and Jeffrey Elman. "Sense and Structure: Meaning as a Determinant of Verb Subcategorization Preferences." *Journal of Memory and Language*, vol. 48, 2003, pp. 281-303.

Ju, Min Kyoung. "Overpassivization Errors by Second Language Learners." *Studies in Second Language Acquisition*, vol. 22, no. 1, 2000, pp. 85-111.

Kaan, Edith. "Predictive Sentence Processing in L1 and L2. What is Different?" *Linguistic Approaches to Bilingualism*, vol. 4, 2014, pp. 257-82.

Lapata, Maria, Frank Keller, and Sabine Schulte im Walde. "Verb Frame Frequency as a Predictor of Verb Bias." *Journal of Psycholinguistic Research*, vol. 30, 2001, pp. 419-35.

Lee, Eun Kyung, Dora Hsin Yi Lu, and Susan Garnsey. "L1 Word Order and Sensitivity to Verb Bias in L2 Processing." *Bilingualism: Language and Cognition*, vol. 16. 2013, pp. 761-75.

MacDonald, Maryellen, Neal Perlmutter, and Mark Seidenberg. "The Lexical Nature

of Syntactic Ambiguity Resolution." *Computer Science Psychological Review*, vol. 101, 1994, pp. 676-703.

Macleod, Catherine, Ralph Grishman, and Adam Meyers. "A Large Syntactic Dictionary for Natural L.anguage Processing." *Computers and the Humanities*, vol. 31, 1997, pp. 459-81.

Marcus, Mitchell, Beatrice Santorini, and Mary Ann Marcinkiewicz. "Building a Large Annotated Corpus of English: The Penn Treebank." *Computational Linguistics*, vol. 19, 1993, pp. 313-30.

Merlo, Paola. "A Corpus-based Analysis of Verb Continuation Frequencies for Syntactic Processing." *Journal of Psycholinguistic Research*, vol. 23, 1994, pp. 435-47.

Montrul, Silvina. "Transitivity Alternations in L2 Acquisition: Toward a Modular View of Transfer." *Studies in Second Language Acquisition*, vol. 22, no. 2, 2001, pp. 229-73.

Oshita, Hiroyuki. "The Unaccusative Trap in Second Language Acquisition." *Studies in Second Language Acquisition*, vol. 23, no. 2, 2001, pp. 279-304.

Pickering, Martin, Matthew J. Traxler, and Matthew W. Crocker. "Ambiguity Resolution in Sentence Processing: Evidence Against Frequency-based Accounts." *Journal of Memory and Language*, vol. 43, 2000, pp. 447-75.

Pickering, Martin, and Roger van Gompel. "Syntactic Parsing." *Handbook of Psycholinguistics*, edited by Matthew J. Traxler and Morton Gernsbacher, Elsevier. 2006, pp. 455-503.

Roland, Douglas, and Dan Jurafsky. "Verb Sense and Verb Subcategorization Probabilities." *The Lexical Basis of Sentence Processing: Formal, Computational, and Experimental Issues*, edited by Paola Merlo and Suzanne Stevenson, John Benjamins, 2002, pp. 325-45.

Trueswell, John, Michael Tanenhaus, and Christopher Kello. "Verb-Specific Constraints in Lexical Preference from Garden-Paths." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 19, no. 3, 1993, pp. 528-53.

국문초록

동사 편향 정보가 제2언어 학습자의 영어 글쓰기에 미치는 영향

김지현
단독 / 한국외국어대학교

본 논문은 동사 편향 정보가 제2언어 학습자의 영어 글쓰기에 미치는 영향에 관하여 연구하였다. 영어를 제2언어로 학습하는 한국어 화자들을 대상으로 글쓰기 과제 실험을 실행하였으며, 43개의 영어 동사를 각각 하나의 영어 문장에 포함하여 쓰게 하였다. 실험자료로 사용한 동사들은 직접목적어 편향 (DO-bias), 보문절 편향 (SC-bias), 그리고 중립편향 (Equibias)으로 나누어졌으며, 실제 각 동사가 직접목적어 또는 보문절 구조에서 사용되는 빈도는 범주형 변수가 아닌 연속변수로, 단일 연속선상에서 선별하였다. 영어 학습자들이 작문한 문장을 분석한 결과, 직접목적어 편향 동사가 가장 많은 직접목적어 구문에 사용되었으며, 보문절 편향 동사에서 가장 적은 수의 직접목적어 구문이 사용되었다. 보문절 구문의 분포는 반대로, 보문절 편향 동사에서 가장 많은 문장이, 직접목적어 편향 동사에서 가장 적은 문장이 사용되었다. 또한, 각 동사별 사용된 직접목적어와 보문절 구문 분포는 기존의 원어민 화자의 코퍼스 데이터와 양의 상관관계를 보임으로써, 본 실험에 참여한 영어 학습자들은 영어 글쓰기에서 동사의 편향 정보를 사용할 뿐만 아니라, 원어민 화자와 같이 편향 정보를 범주형 변수가 아닌 연속변수로 인지한다는 것을 입증하였다.

주제어: 동사 편향, 빈도 기반 정보, 규준화, 제2언어 글쓰기, 영어 학습자

논문접수일: 2021.01.04
심사완료일: 2021.02.09
게재확정일: 2021.02.24

이름: 김지현
소속: 한국외국어대학교 영어통번역학부 부교수
주소: 경기도 용인시 처인구 모현읍 외대로 81
한국외국어대학교 영어통번역학부 (어문관 226호)
이메일: kimjhyon@hufs.ac.kr