Be-Insertion in Interlanguage: A Topic Marker, A Tense/Agreement Morpheme, or Both?

Bora Nam[†]
Indiana University

ABSTRACT

In L2 production, *be*-forms are often inserted before thematic verbs (e.g., *She is study hard*). There have been two hypotheses concerning the function of *be*-forms, that they are topic markers transferred from L1 or tense/agreement morphemes. This study examines the explanatory adequacy of the two hypotheses by comparing acceptability judgments and self-paced listening between Korean and Russian EFL learners. The results support neither hypothesis completely. Contrary to the Topic Marker Hypothesis, both L2 groups accepted *be*-insertion after non-topic as well as topic subjects. Against the T/AGR Morpheme Hypothesis, participants also accepted *be*-insertion after topicalized non-subjects. These results suggest that *be*-forms can have two functions simultaneously, and thus the two hypotheses must not be considered mutually exclusive. The data also call into question the role of L1 transfer, suggesting that the phenomenon may reflect a universal L2 strategy.

Keywords: *Be*-insertion, interlanguage, topic marker, tense/agreement morpheme, L1 transfer

1. Introduction

This study aims to show how the interplay between L1 transfer and UG may result in distinctive form-feature mappings in interlanguage by comparing a particular nontargetlike use of *be*-forms by Korean and Russian EFL learners. In L2 production, a *be*-form is often inserted before a thematic verb, as if it were a particle or a clitic (e.g. She *is* study hard). It is commonly attested in the early stage of L2 English by learners from diverse L1s, such as Arabic (Mourssi, 2013), Bantu languages (Suzman, 1999), Chinese (Yang, 2014), Hmong (Huebner, 1983), Japanese (Shibata, 2006), Korean (Ahn, 2003, 2006; Kim, 2011), Russian (Ionin & Wexler, 2002), Spanish (Fleta, 2003; Mayo et al., 2005), and Turkish (Haznedar, 2007). Not only are the L1s of the learners producing *be*-insertion diverse, but the phenomenon is

[†] Corresponding author: bnam@iu.edu

highly frequent. In Ionin & Wexler (2002), for example, 18 out of 28 Russian child learners produced a sentence containing *be*-insertion in their speaking task. Kim (2011) also showed in his cross-sectional study that the sentences containing *be*-insertion comprised 32% of the written production by Korean learners with low proficiency.

Fleta (2003) observes that the phenomenon is distinct from non-adult-like uses of *be*-forms in child L1 acquisition. The *be*-forms in L1 acquisition substitute for the auxiliary *do* to function as an interrogative particle, as in (1), whereas in L2 acquisition, it is found mostly in declarative sentences, as in (2).

(1) a. *Is* I can do that? b. *Are* this is broke? (Radford, 1990)

(2) Vava *is* want to go to the house. (Suzman, 1999)

The common divergent L2 use of *be*-forms might initially be ascribed to their underspecified meaning in the lexicon as a light verb. *Be*-forms have multiple functions as a copula and as an auxiliary verb in progressive or passive constructions. In most cases, however, there is little content meaning in a *be*-form itself. In this sense, Partee (1999) regards it as a semantically void form, i.e. 'non-lexical *be*', although it may indirectly contribute meaning as a constituent of a construction.

Despite the opaqueness of meaning, *be*-forms are highly frequent in English. According to an analysis of the Oxford English Corpus, they are the second most frequent words following the articles *a(n)* and *the.* Thus, learners are pressured to analyze them from the early stage but have a hard time understanding their constructions and grasping the features associated with them, several of which are strictly functional, such as tense (T) and agreement (AGR). Nontargetlike *be*-insertion can thus be understood as learners' unsuccessful analysis of the structure and/or features associated with the *be*-forms.

There have been two dominant hypotheses concerning the function of *be*-forms in interlanguage: the first suggests that the *be*-forms are a topic marker transferred from the L1 (Ahn, 2003, 2006; Huebner, 1983; Shibata, 2006; Shin, 2001 among others) and the second argues that they are a tense/agreement morpheme (Fleta, 2003; Ionin & Wexler, 2002; Mayo et al., 2005; Yang, 2014 among others). At the base of the Topic Marker Hypothesis lies the premise that there is a dedicated topic morpheme in certain L1s and only learners from such L1s will use *be*-forms to mark topics in their interlanguage. For this reason, I compared topicalization in two L1s,

Korean and Russian, the former with topic markers and the latter without, as well as the target language to make predictions for this study.

Gundel & Fretheim (2004) claim that [+Topic] is universally encoded in languages, but the manner and extent to which it is linguistically encoded differ across languages (e.g. through syntax, prosody, morphology, or combination of these). Table 1 shows that, of these three languages, only in Korean can a topic be marked exclusively with the topic marker *-un/nun,* while word order and sentence stress are the most unmarked means of topic marking in Russian and English, respectively (Schmerling, 1975; Sekerina, 1997).

As the object *ppang/khleb/bread* is topicalized as a contrastive topic, it is followed by the topic marker *-un* in Korean[1]. In Russian, on the other hand, the topic is fronted to the clause-initial position and in English, it gets extra stress and is lengthened in pronunciation. Thus, the Topic Marker Hypothesis predicts that the use of *be*-forms as a topic marker would be limited to Korean learners.

The correspondence between the *be*-forms and the topic markers *-un/nun* in the interlanguage of Korean learners is also supported because the structures in which they are used share a few superficial similarities. Consider the following examples. Just as *-nun* that immediately follows the topic *Na* 'I' in (3), the *be*-auxiliary in (4) is also adjacent to the subject, which is also a topic.

Table 1. Topicalization in Korean, Russian, and English

	Korean	Russian	English
Cannocial order	Jane-i ppang-**ul** coaha-nta Jane-$_{NOM}$ bread-$_{ACC}$ like-PRES	Jane lubit khleb Jane likes bread	Jane likes bread.
Topic-alization	Q: Does Jane like bread and ice-cream?		
	(Jane-i) ppang-**un** coaha-jiman (Jane-$_{NOM}$) bread-$_{TOP}$ like-but	*Khleb* (Jane) lubit no… Bread (Jane) likes but	Jane likes *bread* but...

Moreover, as the form is contracted, it seems closer to the subject/topic than the following verb. Thus, although the *be*-form is generated in V, selecting a VP as its complement in the target grammar, Korean learners are likely to analyze *is* in the input as generated in Top, which can select an IP as its complement, as in (5), due to transfer of the L1 structure in (3). The nonfinite verb *going,* on the other

1) In Korean, subjects and objects can scramble freely, irrespective of the topic, although there is a tendency to prefer the clause-initial position for a topicalized phrase.

hand, can be analyzed as a finite verb as the lexical meaning comes from it, not from the *be*-form. Arguably, learners have not acquired the cues distinguishing finite and nonfinite verbs.

(3) [**TopP** Na_i-**nun** [IP [PredP e_i hakyo-ey ka]-nta]]].
I – TOP school-to go-PRES

(4) [**IP** I'**m**$_i$ [VP t_i [VP going to school]]].

(5) [**TopP** I_i'**m** [IP INFL [PredP e_i go(ing) to school]]]

The learners' analysis also applies to *be*-copulas preceding an AP. As copulas are not required for adjectival predicates in Korean, the *be*-forms preceding an AP can confirm that they are not verbal and thus, they also can be generated in Top.

Supporting the view that *be*-forms are inserted to mark preceding topics, they are often inserted between a topic phrase not governed by the semantics of the finite verb and its sentential comment in the L2 production of Korean and Japanese learners, as in (6-7).

(6) He is friend is many. (Shin, 2001)

(7) Today is my friend come from Kobe. (Shibata, 2006)

On the other hand, the T/AGR Morpheme Hypothesis has been advanced mostly by researchers who have investigated learners from less topic-prominent L1s, such as Arabic, Russian, and Spanish. (Fleta, 2003; Ionin & Wexler, 2002; Mayo et al., 2005; Yang, 2014 among others). Except for a few studies that compared this hypothesis with the Topic Marker Hypothesis (Kim, 2011; Yang, 2002, 2006), most of the studies supporting the T/AGR hypothesis has tried to show that *be*-insertion is neither a progressive nor a passive construction missing surface morphemes. As in (8-9), *be*-forms can be inserted before stative verbs, indicating that they do not denote progressive meaning. As the verbs are also unergative, the preceding *be*-forms are not relevant to the passive construction, either.

(8) He is want go up then. (Ionin & Wexler, 2002)

(9) The boys is no have it. (Fleta, 2003)

The use of *be*-forms as a T/AGR morpheme can also be understood as nontargetlike form-feature mapping. There are two types of forms associated with the T and AGR features in the target language, and the choice is determined by the verbs in the finite position. When the finite verb is a copula or an auxiliary (e.g. *be, have*), the features are mapped onto free morphemes (e.g. *is/are/was/were, has/have/had*) and when the finite verb is a thematic verb, they are mapped onto bound morphemes (e.g. *-s/es, -(e)d*). Between the suppletion and affixal inflection, the former is usually acquired earlier than the latter because it is more efficient (Ionin & Wexler, 2002), and the lexical cues in the former are more salient than the grammatical cues in the latter (Goldschneider & DeKeyser, 2001).

This can cause a problem according to the one-to-one principle (Andersen, 1984). The one-to-one principle explains that due to limitations of learnability and expressibility, there is a preference in the early stage of interlanguage to map one underlying meaning/function to one and only one surface form. When there are multiple forms with the same function, acquiring one form-function relation can inhibit the acquisition of the other forms with the same function. Likewise, when there are two ways to mark agreement and the more efficient one – suppletion – is acquired first, it is likely to be overgeneralized as the only way, preventing the acquisition of the other way – affixal inflection. Consequently, *be*-forms might be analyzed in the interlanguage as a tense/agreement morpheme base-generated in the head of IP and thus, they can precede thematic verbs.

Supporting the T/AGR Morpheme hypothesis, *be*-insertion is most frequent before the acquisition of inflectional affixes and it starts to disappear once these affixes are acquired (Fleta, 2003; Ionin & Wexler, 2002). Yang (2014), who cross-sectionally analyzed Chinese learners' *be*-insertion in written narratives also showed that *be*-insertion was rare in base/nonfinite positions, as in (10), indicating that the *be*-forms are generated in I. Besides, in past contexts, around 50% of *be*-forms were in the past/past participle form, as in (11), from the beginner-level, which also supports that they mark tense in I.

(10) Mrs. Forestier seemed not to be recognize Matilda. (University 1)

(11) The robbery were die. (Secondary 1)

No consensus has been reached regarding these two hypotheses, however, as previous research has relied mostly on production data, which revealed grammatical preference but not acceptability. In the absence of controlling the variables in various contexts, there could be multiple explanations for a particular usage. In addition, none of the studies have directly compared two L1 groups although L1 transfer cannot be claimed without showing inter-L1 group differences, as well as intra-L1 group similarities and inter-L1 group differences (Jarvis, 2000).

In order to address such issues and fill the gap, this study compares the performances of Korean and Russian EFL learners in relation to the following questions.

1. Are the inserted *be*-form analyzed as topic markers in the interlanguages?
2. Are the inserted *be*-form analyzed as agreement morphemes in the interlanguages?

To answer the questions, an acceptability judgment task (AJT) and a self-paced listening task were administered across three experiments.

2. Method

2.1. Participants

In this study, 10 Russian and 11 Korean beginner learners were recruited to test L1 influences in the early stage of interlanguage. As it has been claimed only for the Korean group that the function of *be*-forms changes according to proficiency levels (Kim, 2011), 13 low-intermediate learners were also recruited for the Korean group[2]. Two of the Korean low intermediate participants, however, did not complete all the tasks, so only a part of their data could be included. The age of the participants ranged from 20 to 59. None of the participants had lived in an English-speaking country longer than a month, except for one Korean low-intermediate participant.

2) As an initiative attempt to compare two L2 groups directly concerning the *be*-insertion phenomenon, the main purpose of this study is to test L1 transfer at the beginner's level. As the Korean low-intermediate group cannot be compared with the Russian group, it should be noted that any differences within the Korean group across the proficiency levels cannot be generalized to the Russian group. To pursue more comprehensive understanding of L1 transfer in different proficiency levels, between-L2 group comparison across a wider proficiency range has been advanced in my work-in-progress study.

Table 2. The result of the proficiency test (C-test)

	Russian Beginners	Korean Beginners	Korean Low-intermediate
Mean	14.5/100	13.46/100	41.92/100
s.d.	10.2	6.4	9.2

The participants' general proficiency in English was measured through a C-test. Table 2 shows the group mean scores. The result of a one-way ANOVA showed that there was a significant difference between the three groups, $F = 39.370$, $df = 2$, $p < .001$. The post-hoc analysis revealed that there was no significant difference between the beginner groups but they differed significantly from the Korean low-intermediate group.

2.2. Procedure

Participants signed a consent form, translated in their L1s, and completed a questionnaire before participating in three experiments. The order in which the experiments were given was counter-balanced. Between the experiments, there was at least an hour's break. All the experiments were controlled via DmDX, an online program that enables measurement of reaction times to visual or aural stimuli. After the three experiments, a C-test was given to measure general L2 proficiency. Participants were compensated $15 in their currency after completing all the tasks.

2.3. Tasks and materials

There were four types of items across the three experiments – Subject items in Experiment 1, Object items and PP items in Experiment 2, and INFL items in Experiment 3. Each experiment started with 4 pretest sample items. For each item, the participants read a question, which was shown on the screen for 4 seconds, and pressed a button to listen to the answer segment by segment so that the listening time of each segment could be measured. Then they judged the appropriateness of the answer and checked an option between *appropriate, inappropriate,* and *cannot decide,* which were translated in their L1s.

The stimuli of the processing task were presented aurally as this could minimize the reliance on explicit knowledge and control for the influence of intonation and pronunciation. In each item, there were two clauses with contrastive subjects or

contrastive topics – otherwise, the stimuli would have been unnatural as the other kinds of topics that can be recovered from contexts are usually dropped in their interlanguage, as well as in their L1s. Besides, by using conjoined clauses, the number of measurements could be doubled. The processing task was expected to detect more subtle differences in judgment. A slow-down in processing can indicate that the given stimulus did not conform to what had been expected by the participants as the structure was disallowed in their interlanguage grammar.

The Subject, Object, and PP items were created to examine the judgment of *be*-forms as a topic marker. There were a total of 16 Subject items, which could be divided into four types depending on the topicality of the subjects and the presence/absence of *be*-insertion. Topics in this study were defined as those that are present in the common ground of the interlocutors and are likely to be present in the immediate common ground (Zimmermann & Féry, 2010). As illustrated in (12), the topicality was manipulated by the preceding questions. The topic question '*What do your mom and dad do…?*' can bring the other interlocutor's mom and dad into the common ground and thus they become the topics in the following answer, whereas the Nontopic question '*Why can't you sleep well at night?'* cannot do so[3].

(12) **Topic Q:** What do your mom and dad do during the weekend?
Nontopic Q: Why can't you sleep well at night?
Answer in segments[4]

	S1	V1	XP1	Conj	S2	V2	XP2
+be	My mom is	play	the guitar	and	my dad is	sing	songs
−be	My mom	plays	the guitar	and	my dad	sings	songs

To compare processing time across the item types, the length of each segment in the answers was controlled – as the number of syllables in a segment could differ across conditions (e.g. *My mom is* vs. *My mom*), there was a longer pause before the shorter segments were played.

3) A reviewer pointed out that the nontopic question can induce topicalization in the following answer (e.g. *The reason I could not sleep well*). The manipulation of topicality in this study, however, focuses on the phrases immediately preceding *be*-forms – in this example, the hearer's mom and dad – and distinguishes whether they are present in the common ground or not; as, to my understanding, the Topic Marker Hypothesis claims that *be*-forms mark the topicality of the preceding phrases rather than implicit concepts.

4) When the given question started with *why,* the phrase *because* was added at the beginning of the answer but due to a presentation issue, the phrase is missing in the sample.

If the *be*-forms were analyzed as topic markers, the *be*-insertion after nontopic subjects would be accepted at a significantly lower rate compared to the *be*-insertion after topic subjects, and a slow-down in processing would be expected only for the former – as this could not be expected in the interlanguage grammar. If the *be*-forms were analyzed as T/AGR morphemes, on the other hand, the judgment of *be*-insertion would not be affected by the topicality of the subjects and there would be no slow-down in processing the *be*-insertion after nontopic subjects.

In Experiment 2, there were 8 Object items and 8 PP items, each of which could be divided into two types depending on the presence/absence of *be*-insertion. As *be*-forms are usually inserted after the initial phrases, all the objects and PPs in these items had to be topicalized and fronted to precede a *be*-form, as in (13-14) and thus, the influence of topicality could not be tested.

(13) **Q for Object items:** Are you good at Japanese and Chinese?
Answer in segments

	O1	S1	V1	XP1		O2	S2	V2	XP2
+be	Japanese is	I	understand	a little	and	Chinese is	I	speak	very well
−be	Japanese	I	understand	a little	and	Chinese	I	speak	very well

(14) **Q for PP items:** What do you do on Saturday and Sunday?
Answer in segments

	PP1	S1	V1	XP1		PP2	S2	V2	XP2
+be	On Saturday is	I	play	games	and	on Sunday is	I	sing	songs
−be	On Saturday	I	play	games	and	on Sunday	I	sing	songs

The Topic Marker Hypothesis predicts that the Object and PP items containing *be*-insertions and their targetlike counterparts should be accepted at a similar rate without any slow-down in processing. The T/AGR Morpheme Hypothesis, on the other hand, predicts that the *be*-insertion after the topic but nonsubject phrases should not be accepted, resulting in a slow-down in processing as the *be*-forms were not in the head of IP and they did not agree with the subjects but with the preceding topics.

Finally, there were 16 INFL items[5] in Experiment 3 to test the judgment of

5) Judgment of the items does not necessarily require T/AGR split in the interlanguage so they were termed INFL items.

be-forms as an agreement morpheme, as well as a topic marker. Unlike the Subject items, in all the INFL items, the subjects were followed by a *be*-form, and the items were divided into four types depending on the inflection on the following verbs and the topicality of the subjects, as in (15).

(15) **Topic Q:** Why do your brother and sister want to go to China?
Q: Why are your parents interested in China?
Answer in segments

		S1	V1	XP1		S2	V2	XP2
−Infl	Because	my brother is	study	Chinese	and	my sister is	like	pandas
+Infl	Because	my brother is	studies	Chinese	and	my sister is	likes	pandas

The Topic Marker Hypothesis predicts that the judgment of the INFL items would be affected greatly by the topicality of the preceding subjects but the T/AGR Morpheme Hypothesis predicts that regardless of the topicality of the subjects, the *be*-forms followed by inflected verbs would be accepted at a low rate – if *be*-forms were analyzed as an agreement morpheme, the inflection on the following verbs indicates that agreement was marked redundantly in the clause[6].

Table 3. The composition of items across the experiments

		Experimental items		Distractors
Exp. 1 (40 items)	Subj items	Nontopic S + V (4)	Topic S + V (4)	Grammatical (12)
		Nontopic S + *be* +V (4)	Topic S + *be* +V (4)	Ungrammatical (12)
Exp. 2 (40 items)	Object/ PP items	O + SV (4)	PP + SV (4)	Grammatical (12)
		O + *be* +SV (4)	PP + *be* +SV (4)	Ungrammatical (12)
Exp. 3 (56 items)	INFL items	Nontopic S + *be* + V (4)	Topic S + *be* + V (4)	Grammatical (20)
		Nontopic S + *be* + V*es* (4)	Topic S + *be* + V*es* (4)	Ungrammatical (10)

In order to conceal the purpose of the experiment and to minimize the influence of explicit knowledge, there were also distractors in each experiment. In Experiments 1 and 2, half of the distractors were grammatical, and the other half were ungrammatical. In Experiment 3, on the other hand, two-thirds of the distractors

6) This prediction is based on the assumption that be-forms occupying the INFL node block agreement-checking with the inflected thematic verbs at PF.

were grammatical as all the INFL items were not grammatical in the target language. Table 3 summarizes the composition of items in Experiments 1-3.

3. Results

The influence of the manipulated factors – topicality of the preceding subjects, presence/absence of *be*-insertion, and inflection of the following verbs – on the acceptability judgment results were examined through a mixed Analysis of Variance (ANOVA). The listening times, on the other hand, were analyzed through a Mixed Linear Model. Before running the analysis, the outlier listening times which exceeded the mean time plus two standard deviations were replaced with the mean value, which changed 7% of the data. The refined listening times were then transformed to log 10 values and the data from each group were analyzed separately – as the general listening time differed greatly across groups, regardless of the conditions.

3.1. Subject items

Figure 1 shows the mean acceptance rates of the Subject items for each group. A mixed ANOVA revealed significant main effects of *Be*-insertion ($F(1, 30) = 2.171$, $p = .001$, partial $\eta^2 = .320$) and Topicality ($F(1, 30) = .107$, $p = .054$, partial $\eta^2 = .118$). The results show that regardless of the topicality of the subjects, *be*-insertion significantly lowered the acceptance rates and so did nontopic subjects regardless of whether *be*-forms were inserted or not.

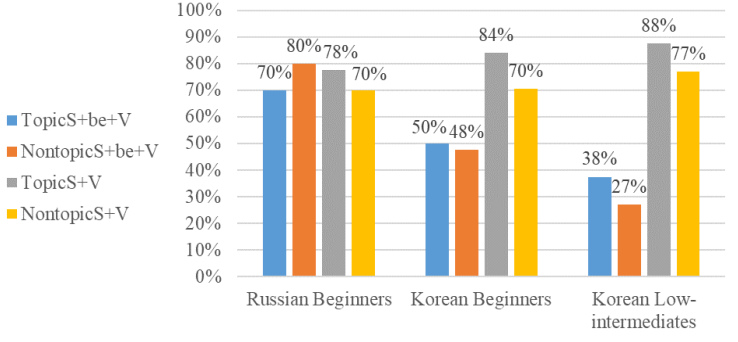

Figure 1. Mean acceptance rates of Subject items

In addition, there were a significant interaction of *Be*-insertion*Group ($F(2, 30) = 1.435$, $p = .017$, partial $\eta^2 = .237$) and a trend of interaction of Topicality * *Be*-insertion ($F(1, 30) = .076$, $p = .081$, partial $\eta^2 = .098$). As the left graph in Figure 2 shows, only the Korean low-intermediate group could distinguish the *be*-insertion items from their targetlike counterparts while the Korean and Russian beginner groups were not sensitive to the nontargetlike insertion of *be*-forms.

The right graph in Figure 2, on the other hand, shows that the effect of Subject Topicality tended to be found only in the targetlike items, although there was overlap in the 95% confidence interval (CI), indicating that the difference was not significant.

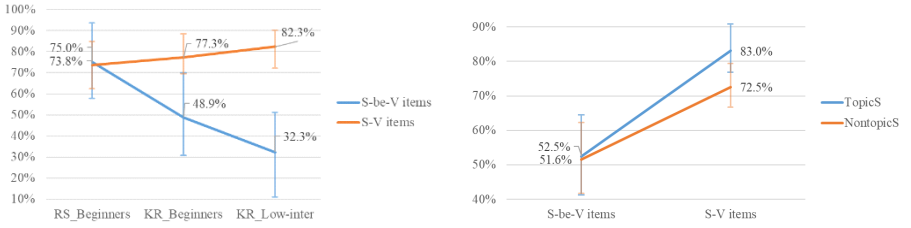

Figure 2. Interactions of groups*Be-insertion and Topicality*Be-insertion in the results of Subject items

To summarize the AJT results of Subject items, *be*-insertion, as well as nontopic subjects, tended to lower the overall acceptance rates, but the degree of influence differed across groups – only the Korean low-intermediate accepted the *be*-insertion items at a significantly lower rate that the targetlike counterparts and between the beginner groups, the Korean group was more sensitive to the nontargetlike *be*-insertion. Unlike the prediction for the Topic Marker Hypothesis, however, the acceptance rate of *be*-insertion did not differ significantly depending on the topicality of the subjects.

The processing of each group, on the other hand, was analyzed separately through a Mixed Linear Model in SPSS with fixed effects of Topicality (topic subject, nontopic subject) and *Be*-insertion (*be*-insertion, targetlike), and random effects of Subject and Question number. As Table 4 shows, *be*-insertion significantly slowed down the processing of all the groups in the S region. It also slowed down the processing of the Korean groups in the V regions. The processing of the Russian group in the region, on the other hand, significantly slowed down after nontopic subjects, regardless of presence/absence of *be*-insertion, and the effect of *be*-insertion was limited to topic subjects. The overall results indicate that the insertion of

be-forms was not expected in the interlanguage of any groups regardless of the topicality of the preceding subjects and thus, the participants had to slow-down their processing to reanalyze the unexpected stimuli.

Table 4. Processing of Subject items (A > B: Slow-down in condition A)

	S	V	XP
	My mom My mom is	goes go	to work
RS-B	*Be*-insertion > Targetlike	Effect 1 Nontopic S > TopicS Effect 2 Nontopic S → *Be*-insertion = Targetlike Topic S → *Be*-insertion > Targetlike	n.s.
KR-B KR-LI		*Be*-insertion > Targetlike	

3.2. Object and PP items

Figure 3 shows the mean acceptance rates of the Object and PP items for each group. In the results of a mixed ANOVA for the Object items, neither *Be*-insertion nor interaction of *Be*-insertion*Group was significant. The results for the PP items, on the other hand, revealed a significant main effect of *Be*-insertion ($F(1, 31) = .714$, $p = .005$, partial $\eta^2 = .225$), indicating that *be*-insertion items were accepted at a significantly lower rate than the targetlike counterparts. There was also a significant interaction of *Be*-insertion * Group ($F(2, 31) = .460$, $p = .070$, partial $\eta^2 = .157$).

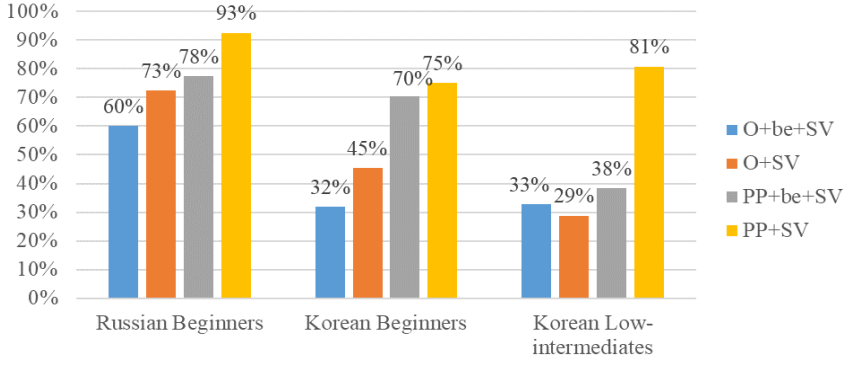

Figure 3. Mean acceptance rates of Object and PP items

The left graph in Figure 4 shows that Russian groups accepted the Object items more than the Korean groups, while none of the groups distinguished the presence/absence of *be*-insertion after the fronted objects. The right graph, on the other hand, shows that the Korean and Russian beginner groups accepted the 'PP-*be*-SV' items and 'PP-SV' items at similar rates and only the Korean low-intermediate group accepted the *be*- insertion items significantly less than their targetlike counterparts.

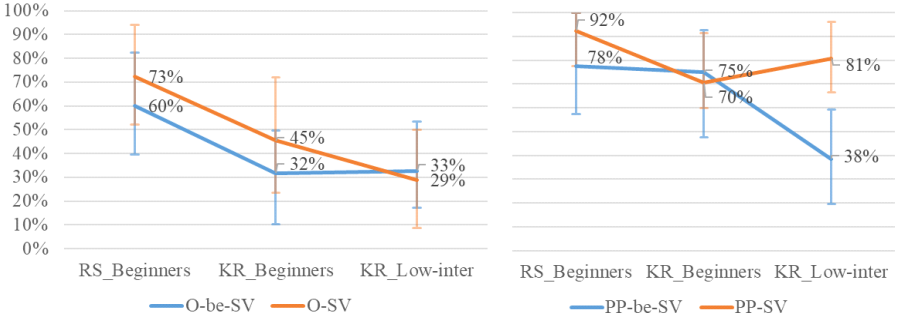

Figure 4. Interaction of Be-insertion*Group in the results of Object and PP items

The AJT results for the Object and PP items suggest that the Russian group was not sensitive to the insertion of *be*-forms after topic objects and PPs. The judgment of *be*-insertion after topic objects by the Korean groups, however, could not be examined as they rejected object fronting regardless of whether the objects were followed by a *be*-form or not. *Be*-insertion after topic PPs, on the other hand, tended to be accepted by the Korean beginner group and rejected by the low-intermediate group.

The processing of the Object and PP items by each group was also analyzed through a Mixed Linear Model in SPSS with a fixed effect of *Be*-insertion (*be*-insertion, targetlike), and random effects of Subject and Question number. As Table 5 shows,

Table 5. Processing of Object items (A > B: Slow-down in condition A)

	O	S	V	XP
	Vegetable Vegatable is	I	eat	sometimes
RS-B	Targetlike > *Be*-insertion	n.s.	n.s.	n.s.
KR-B	*Be*-insertion > Targetlike			
KR-LI	n.s.		*Be*-insertion > Targetlike	

the insertion of *be*-forms after topic objects significantly facilitated the processing of the Russian group in the O region while it inhibited the processing of the Korean groups – the beginner group slowed down in the O region and the low-intermediate group, whose processing rate was faster, slowed down in the V region due to a spill-over effect.

Table 6, on the other hand, shows that *be*-insertion after topic PPs did not affect the processing of the Russian beginner and Korean low-intermediate groups. The processing of the Korean beginner group, however, was at first inhibited in the PP region and facilitated in the V and XP regions, indicating that the *be*-forms after PPs were at first unexpected, but they were analyzed immediately and could help the processing afterward.

Table 6. Processing of PP items (A > B: Slow-down in condition A)

	PP	S	V	XP
	On Saturday On Saturday is	I	play	games
RS-B	n.s.		n.s.	n.s.
KR-B	*Be*-insertion > Targetlike	n.s.	Targetlike > *Be*-insertion	Targetlike > *Be*-insertion
KR-LI	n.s.		n.s.	n.s.

The overall processing results show that unlike *be*-insertion after subjects, *be*-insertion after fronted topics usually did not affect processing, indicating this type of *be*-insertion was allowed in the interlanguage. Moreover, the *be*-forms after topic objects even facilitated the processing of the Russian group in the Object region and those after topic PPs facilitated the processing of the Korean beginner group in the V and XP regions. These findings suggest that the insertion of *be*-forms tended to be integrated better after topicalized nonsubjects than subjects in the interlanguage of the Russian and Korean participants.

3.3. INFL items

Figure 5 shows the mean acceptance rates of the INFL items. The results of a mixed ANOVA showed significant main effects of Topicality ($F(1, 32) = .656$, $p = .002$, partial $\eta^2 = .266$) and Inflection ($F(1, 32) = .290$, $p = .031$, partial $\eta^2 = .138$) – across groups, *be*-insertion was accepted at a significantly higher rate when the

preceding subjects were topics than nontopics and the following thematic verbs were inflected rather than uninflected.

There was also a significant interaction of Topicality*Group ($F(2, 32) = .516$, $p = .018$, partial $\eta^2 = .222$). As Figure 6 shows, the judgments of the beginner groups were affected greatly by the topicality of the preceding subjects while the Korean low-intermediate group tended to reject *be*-insertion regardless of the topicality.

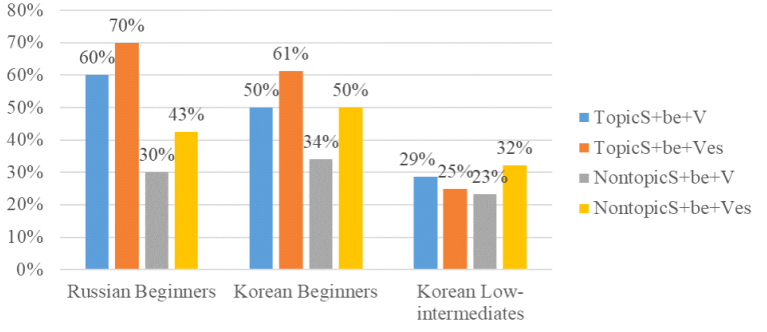

Figure 5. Mean acceptance rates of INFL items

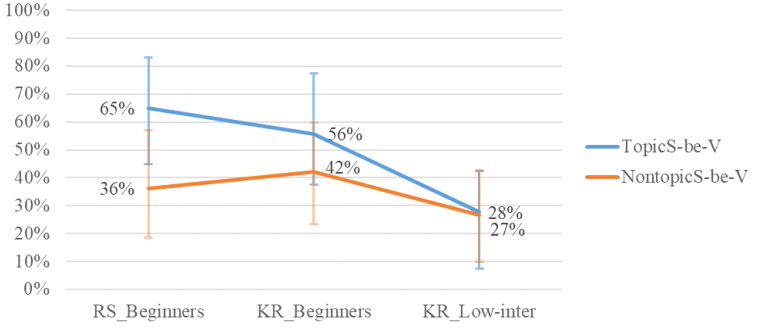

Figure 6. Interaction of Topicality*Group in the results of INFL items

As the processing of the INFL items was analyzed through a Mixed Linear Model with fixed effects of Topicality (topic subject, nontopic subject) and Inflection (inflected, bare), and random effects of Subject and Question number, the results also confirmed the effect of Topicality. As Table 7 shows, when the subjects before *be*-insertion were topics, the processing of the Russian group was facilitated in the S region and that of the Korean groups in the V region. The processing of the Russian group was also facilitated in the XP region by the inflection on the thematic

verbs only when the subjects were nontopics. Unlike the AJT results, however, there was no significant main effect of Inflection in the processing results.

Table 7. Processing of INFL items (A > B: Slow-down in condition A)

	S	V	XP
	Bill is	buy buys	a cake
RS-B	Nontopic S > Topic S	n.s.	Nontopic S → bare > inflected Topic S → inflected = bare
KR-B KR-LI	n.s.	Nontopic S > Topic S	n.s.

4. Discussion

Q1. Are the inserted be-forms analyzed as topic markers?

The Topic Marker Hypothesis predicted that only the Korean groups would analyze the inserted *be*-forms as a topic marker due to the transfer of the topic marker *-un/nun* from the L1. While the Korean low-intermediate group tended to reject all kinds of *be*-insertion, indicating their interlanguage was gradually converging on the target grammar as proficiency increased, the results of the Korean beginner group partially supported the predictions. Their judgment of the PP items showed that they accepted the *be*-insertion after topicalized PPs. Moreover, *be*-insertion after PPs also facilitated their processing in the V and XP regions. These results can be ascribed to their analysis of *be*-forms as a topic marker.

Against the predictions for the hypothesis, however, Both the Korean beginner and the low-intermediate groups accepted *be*-insertion after nontopic subjects at a similar rate as the one after topic subjects, questioning the view that such *be*-forms could function only as topic markers in their interlanguage. Although topic subjects increased the acceptance rate of the INFL items, all of which contained *be*-insertions, it was unclear either whether the rate increased as the *be*-forms were analyzed as a topic marker or simply because topic items tended to be accepted more than nontopic items – considering that topic items were accepted at a higher rate than nontopic items among the Subject items, regardless of the presence or absence of *be*-insertion. The processing of Subject items also casts doubt on the hypothesis, as

topic subjects did not facilitate the processing of *be*-insertion items.

There were also results from the Russian group that could not be explained by the Topic Marker Hypothesis. One prediction was that it was unlikely for the Russian participants to analyze inserted *be*-forms as topic markers as there is no morpheme designated to topic marking in Russian. The Russian group, however, was not sensitive to the insertion of *be*-forms above I in the Object and PP items. In the results of the INFL items, the topicality of the preceding subjects also affected the Russian judgments more greatly than those of the Korean groups. Moreover, the processing of PP items was not inhibited by *be*-insertion and the processing of the Object items was even facilitated in the Object region. These results support the idea that the Korean learners were analyzing the *be*-forms as topic markers, although the analysis could not result from L1 transfer.

To summarize, the results of the Korean beginner group concerning the PP items suggest that they mayd analyze the inserted *be*-forms as topic markers. Their results of the Subject and INFL items, however, suggest that the function of the *be*-forms may not be limited to topic marking as the judgment and processing of *be*-insertion after subjects were not greatly affected by the topicality of the preceding subjects. Besides, questioning the view that the analysis of *be*-forms as a topic marker results from L1 transfer, the Russian group also tended to accept *be*-insertion after topicalized nonsubjects without slowing down processing.

Q2. Are the inserted be-forms analyzed as agreement morphemes?

The T/AGR Morpheme Hypothesis predicted that both the Korean and Russian groups could analyze the inserted *be*-forms T/AGR morphemes. Consistent with this prediction, their judgment and processing of *be*-insertion after subjects were not affected by the topicality of the subjects. As explained above, however, they tended to accept *be*-insertion after topicalized nonsubjects, although the *be*-forms were inserted above I and did not agree with the subjects but with the topicalized phrases in the items.

The results of the INFL items, on the other hand, are more complicated to interpret. At first, I predicted that if the *be*-forms were analyzed a T/AGR morphemes, the acceptance rate would be lowered when the following verbs were inflected as redundant agreement marking may not be allowed in the interlanguage. If the *be*-forms were analyzed as topic markers, on the other hand, the acceptability of *be*-insertion would not differ depending on the inflection on the following verbs, considering that the Korean and Russian beginner groups were not sensitive to the

inflection on thematic verbs (e.g. *She like* vs. *She likes*) in the distractor items.

Conforming neither to the former nor the latter predictions, however, the AJT results of the INFL items showed that across all groups, the acceptance rate increased when the *be*-forms were followed by inflected verbs rather than uninflected ones. In other words, *be*-insertion after subjects raised the sensitivity to agreement marking with the inflectional morpheme *-s*.

I suggest that this result is more compatible with the T/AGR Morpheme Hypothesis than the Topic Marker Hypothesis on the assumption that agreement can be encoded with multiple means at one time in the interlanguage. Supporting that such a learner analysis is available through UG, in several dialects of Dutch, agreement can be marked via multiple means at one time, including nonverbal elements, such as clitics, *pro*s, and complementizers, as well as verbal inflection (de Vogelaer, 2003). As in (16), in addition to the verbal inflection *-en,* the 3rd plural agreement can also be marked with the clitic *ze* and the weak pronoun *zieder.* In (17), the agreement is marked with the complementizer *dan,* as well as the verbal inflection.

(16) Ze hebb-en (zieder) drie kinderen.
They.cl have.3pl (they.pro) three kids
'They have three kids.'

(17) Weten ze da**n**-ze moeten komen? Jaa**n**-s
Know.3pl they that.3pl-they must come Yes.3pl-they
'Do they know that they must come? Yes (they) [do]'
(de Vogelaer, 2003)

Given that the interlanguage of the Korean and Russian participants could allow redundant agreement marking, their results of the INFL items are consistent with the view that the *be*-forms were analyzed as optional agreement morphemes. As lexical morphemes are relatively salient, the insertion of *be*-forms as agreement morphemes could raise the participants' sensitivity to overt agreement marking, including the inflectional agreement on the following verbs. Thus, they could distinguish whether agreement was correctly marked with the inflectional morpheme *-s* or not only when a *be*-form was inserted before the verb.

To summarize, the nonsignificant difference in the acceptance rate of the *be*-insertion items depending on the topicality of the preceding subjects was expected

by the T/AGR Morpheme hypothesis but not by the Topic Marker Hypothesis. More directly supporting the former, the results of the INFL items also showed that overt agreement marking on the thematic verbs could be distinguished better when *be*-forms were inserted before the verbs. With the assumption that the Korean and Russian participants allowed redundant agreement marking via multiple means, the evidence suggests that *be*-forms can be optional agreement morphemes which could raise the sensitivity to overt agreement marking with verbal inflection. The function of *be*-forms, however, did not seem to be limited to agreement morphemes as the *be*-forms inserted after topicalized nonsubjects were also accepted without any slow-down in processing even when they did not agree with the subjects.

5. Conclusion

This study yielded mixed results for both L2 beginner groups. The nonsignificant effect of Topicality in processing and judging *be*-insertion suggests that *be*-forms may not function solely as topic markers in interlanguage. Acceptance of *be*-forms after topicalized/fronted PPs and objects, on the other hand, suggests that *be*-forms may not be placed in the I node and thus may neither function solely as a T/AGR morpheme. The latter result, however, is also incompatible with the Topic Marker Hypothesis as the hypothesis predicted distinctive performances between the Korean and Russian L2 groups due to distinctive influences from the L1s.

These findings call for two revisions of the previous research that investigated the *be*-insertion phenomenon. First, the Topic Marker Hypothesis and the T/AGR Morpheme Hypothesis can no longer be treated as mutually exclusive, as neither the former nor the latter alone can explain the results completely – the two hypotheses claim that the *be*-forms could mark either topics or tense/agreement but not both, but the results reveal that they can support both functions in the interlanguages of the L2 groups.

Second, the source of the interlanguage grammar is not limited to L1 transfer. The Topic Marker Hypothesis claims that the nontargetlike analysis of the *be*-forms as topic markers reflects L1 transfer, but some results contradict this account. The Russian beginners, as well as the Korean beginners, appear to analyze the *be*-forms as topic markers, although there is no morpheme dedicated to topic marking in Russian. Likewise, the Korean group seem to analyze the *be*-forms as agreement morphemes even though agreement is not overtly marked in the L1. The similar

analyses of the *be*-forms by the L2 groups despite differences in the L1s suggest that the *be*-insertion phenomenon may reflect a universal L2 strategies, indicative of UG active in L2 acquisition.

Instead of revising the Topic Marker Hypothesis or the T/AGR Morpheme Hypothesis, one possibility might be to abandon both hypotheses and pursue alternative accounts. To begin with, there may be a chance that the inserted *be*-forms were not noticed in the aural stimuli and thus were not analyzed at all. This possibility, however, is unlikely given by the processing results of the Subject items – *be*-insertion resulted in a significant slow-down in processing compared to the targetlike counterparts, which indicates that the structure was noticed. Alternatively, another possibility could be that the inserted *be*-forms are neither topic markers nor T/AGR morphemes but something else, similar to second position clitics in classical Greek, pragmatic fillers (e.g. *well, umm*) or phrase boundary markers. Such alternatives, however, have not been discussed seriously in a theoretical framework, so they remain unmotivated.

In conclusion, this study suggests that the Topic Marker Hypothesis and the T/AGR Morpheme hypothesis are not mutually exclusive. However, the study also has limitations that need to be addressed in future work. First, this study relied only on an acceptability judgment task, but a more comprehensive understanding of this phenomenon can be reached only by triangulating the results with production data. In addition, in this study, testing two proficiency levels only for the Korean group made the reader expect contradict with the Russian group. In order to show precisely how the function of *be*-forms changes according to the L1 and proficiency levels, however, learners from a wider range of proficiency need to be tested for both groups. Moreover, this study examined the acceptability of *be*-forms used as agreement morphemes, but those used as tense morphemes need to be investigated, as well, to test more fully the explanatory adequacy of the T/AGR Morpheme Hypothesis. Finally, a control group – native speakers of English – must be included to enhance the validity of the tasks, as their results would confirm that the tasks successfully measured what they were supposed to. A more comprehensive understanding of *be*-insertion will contribute to more fine-tuned understandings of how form-feature mapping changes in the development of interlanguage and of how each stage in the process can be conditioned by the interplay between L1 transfer and UG.

Abbreviations

3pl = third person plural, ACC = accusative, Cl = clitic, I/INFL = inflection, NOM = nominative, PRES = present, TOP = topic

References

Ahn, S. H. (2003). A note on the topic-comment stage in Korean EFL syntactic development. *Studies in Generative Grammar, 13*(2), 369-32.

Ahn, S. H. (2006). The grammar of verb be in early Korean EFL interlanguages. *Studies in Generative Grammar*, *16*(4), 769-782.

Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language learning*, *34*(4), 77-95.

de Vogelaer, G. (2003). Person marking in Dutch dialects. In B. Kortmann (Ed.), *Dialectology meets typology: Dialect grammar from a cross-linguistic perspective* (pp 181-210). Berlin: Mouton de Gruyter.

Fleta, M. T. (2003). Is-insertion in L2 grammars of English: A step forward between developmental stages. In J. M. Liceras, H. Zobl, & H. Goodluck (Eds.), *Proceedings of the 6th generative approaches to second language acquisition conference* (GASLA, 2002) (pp. 85-96). Somerville, MA: Cascadilla Proceedings Project.

Goldschneider, J. M. and DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 51*(1), 1-50.

Gundel, J. K. and Fretheim, T. (2004). Topic and focus. In L. Horn, & G. Ward (Eds.), *The handbook of pragmatics* (pp 175-196). Oxford: Blackwell.

Haznedar, B. (2007). The acquisition of tense—aspect in child second language English. *Second Language Research*, *23*(4), 383-417.

Huebner, T. (1983). *A longitudinal Analysis of the Acquisition of English.* Ann Arbor, MI: Karoma Publishers.

Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'?: Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research, 18*(2), 95-136

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence in the interlanguage lexicon. *Language Learning*, *50*(2), 245-309.

Kim, K. (2011). Overgenerated *be* from topic marker to verbal inflection. *Foreign Language Education Research, 14,* 1-22.

Mayo, G., del Pilar, M., Ibarrola, A. L., & Liceras, J. M. (2005). Placeholders in the English interlanguage of bilingual (Basque/Spanish) children. *Language Learning, 55*(3), 445-489.

Moursi, A. (2013). Crosslinguistic influence of L1 (Arabic) in acquiring linguistic items of L2 (English): An empirical study in the context of arab learners of English as undergraduate learners. *Theory & Practice in Language Studies*, *3*(3), 397-403.

Partee, B. H. (1999). Focus, quantification, and semantics-pragmatics issues. In P. Bosch, & R. van der Sandt (Eds.) *Focus, Linguistic, Cognitive, and Computational Perspectives* (pp. 213-231). Cambridge: Cambridge University Press.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell.

Schmerling, S. (1975). Evidence from sentence stress for the notions of topic and comment. *Texas Linguistics Forum, 19,* 135-141.

Sekerina, I. A. (1997). *The syntax and processing of scrambling constructions in Russian* (Doctoral dissertation). City University of New York.

Shibata, M. (2006). Topic marking in English composition by Japanese EFL learners. *Scripsimus, 15,* 1-26.

Shin, J-S. (2001). L1 influence in foreign language learning: Topic-prominence in Korean EFL learners' interlanguage grammar. *Foreign language education*, *8*(1), 1-21.

Suzman, S. (1999) The acquisition of English as a second language by children who speak Bantu languages. Paper presented at the 8th International Congress for the Study of Child Language, San Sebastian, 1999.

Yang, H. K. (2002). Korean EFL learner's acquisition of English inflectional features. *Korean Journal of English Language and Linguistics, 2*(2), 227-248.

Yang, H. K. (2006). Functional projection in Korean ESL learner's interlanguage syntax. *The SNU Journal of Education Research, 15,* 117-128.

Yang, S. (2014) The functions of the nontarget be in the written interlanguage of Chinese learners of English, *Language Acquisition, 21*(3), 279-303.

Zimmermann, M., & Féry, C. (2010). *Information structure: Theoretical, typological, and experimental perspectives*. Oxford University Press.

Bora Nam
PhD Candidate
Department of Second Language Studies
Indiana University
1165 E. Third Street, Bloomington, Indiana, United States
E-mail: bnam@iu.edu

Received: October 2, 2019
Revised version received: November 5, 2019
Accepted: December 18, 2019