Why *IS* 'be' appear there?: Topic marker vs. underdeveloped functional category*

Incheol Choi

(Kyungpook National University)

Choi, Incheol. 2013. Why *IS* 'be' appear there?: Topic marker vs. underdeveloped functional category. *Linguistic Research* 30(3), 567-581. Korean learners who learn English as a foreign language tend to use the verb *be* excessively in the initial stage of acquisition. Traditionally, there have been two opposing explanations about these errors. The first one is "topic marker view" (Ahn, 2003) and the other is "functional category view" (Yang, 2001). I carried out a corpus study using a learner corpus called KELC to examine which view explains the phenomenon more plausibly. After checking 1,235 error sentences I conclude that the errors are not explained by a sole reason. Instead, I suggest that the errors involve both the language transfer effect that is due to the topic prominence of Korean and the underdeveloped functional categories appearing in learners' developmental stages. **(Kyungpook National University)**

Keywords *be*-insertion errors, topic marker, underdeveloped functional category, English learner corpus

1. Introduction

Korean learners who learn English as a foreign language tend to use the verb *be* excessively in the early stage of acquisition. In many cases, such excessive use of the *be*-form results in ungrammatical sentences as shown in (1).

(1) a. *The girl is very very like dog.
   b. *They were arrived at the pizza store.
   c. *He is eat hamburger.

The sentences in (1) are erroneous in that the *be*-forms are neither progressive nor passive auxiliaries. This type of *be*-insertion errors are prevalent among Korean learners of English in the early stage of acquisition and persist until the learners advance to a highly advanced level.

Traditionally, there have been two opposing arguments concerning Korean learners' erroneous use of *be* verbs. The first one is "topic marker view" (Hahn, 2000; Shin, 2000; Ahn, 2003) and the other is "functional category view" (Yang, 2001). In the former view, the erroneously inserted *be* is considered a topic marker, which is a characteristic of Korean grammar. That is, the excessive use of the *be*-form by Korean learners is due to the language transfer effect. On the other hand, the latter view defines the inserted *be* as an underdeveloped verbal functional category. In this paper, I investigate the two possibilities and suggest that both views have empirical supports. To do that, I carry out a corpus study by using KELC (KNU English Learner Corpus), which consists of about 160,000 words from 838 written samples of elementary school students.

2. Literature review

In this section, I will review the two previous approaches on the cause of Korean learners' *be* insertion errors. First, the topic marker view proposes that *be* plays a role as a topic marker because of L1 (first language) transfer (Hahn, 2000; Shin 2000, 2001; Ahn, 2003). Korean is generally assumed to have both characteristics of the topic oriented language and the subject oriented language in that it has the topic in addition to the subject (cf. Li and Thomson, 1976). The topic in Korean does not form a selectional relation with the head verb, and it announces the theme of the discourse as the center of attention of a sentence.

(2) kim-un   apeci-ka   puca-i-si-ta
    Kim-TOP   father-NOM   rich-COP-DECL
    'As for Kim, his father is rich.'

The initial NP *Kim-un* in (2) is not thematically selected by the head predicate *puca-i-ta* and, therefore, does not establish the honorific agreement relation with the

predicate. Instead, the NP is an entity that is described by the following part of the sentence. Shin (2001) suggests that this characteristic of the topic language has an impact on Korean learners' acquisition of English. Most conspicuous effect is the excessive use of the *be*-form. According to Shin (2001), the erroneous *be* as in (1) is in fact a topic marker like −*nun* in Korean (see also Hahn, 2000). Ahn (2003) also argues that the erroneously inserted *be* is a topic marker and cannot be analyzed as an inflectional category. For example, there are many cases in which the inserted *be* inflects for agreement but not for tenses, as shown in (3).

(3) These problems are disappeared. (Ahn, 2003)

The sentence in (3) describes a past event but the tense is marked only on the second main verb *disappear*, not on the *be*-form *are*.

Second, negation does not attach to the *be*-form but in many cases a negative auxiliary word *don't* appears between *be*-forms and the following main verbs as shown in (4).

(4) a. We are don't do anything.
    b. He is don't have girlfriends. (Ahn, 2003)

Ahn (2003) argues that these distributional properties show that the erroneous *be*-form cannot be analyzed as an inflectional category. If the *be*-form was an inflectional functional category, learners would not need additional auxiliary word after it.

On the other hand, Yang (2001) suggested that the erroneous *be*-form in Korean learners' interlanguage is an underdeveloped verbal functional category. Yang (2001) observes that the *be*-form produced by Korean learners functions as agreement and tense.

(5) a. Sally is eat bread.
       'Sally eats bread.'
    b. mummy is make breakfast.
       'Mummy makes breakfast.'

    c. she is egg eat?
      'Does she eat eggs?'
(6) a. I'm like soccer.
      'I like soccer.'
    b. He is live Lucky Apartment.
      'He lives in Lucky Apartment.'
    c. Danny is play computer.
      'Danny plays a computer game.'

In (5) and (6), the main verbs do not agree with their subject with respect to tense and agreement but the following *be*-forms do. Yang (2001) takes this observation as supporting evidence for his underdeveloped verbal functional category hypothesis. Yang further suggests some counterevidence against the topic marker views.

(7) a. He, family is four.
      'As for him, he has four members in his family.'
    b. He, family is father, mother, and Jung-Jae
      'As for him, his family consists of father, mother, and
      Jung-Jae.'
    c. She, subject is math
      'As for her, her favorite subject is math.'

Yang (2001) points out that when two NPs appear in a Korean sentence, such as "NP+NP+*be*+X" structure, *be* follows the second NP, not the first NP. Given that the first NP is construed as a topic, the topic marker view would expect the first NP to be followed by the *be*-verb. However, Yang argues that such sentences are not found in utterances of learners who are in early stages of learning.

    Scholars suggest that the excessive use of *be* is also observed in L2 learners of other L1 backgrounds (Zobl, 1989; Yip, 1995). They often point out unaccusative verbs as the main cause of the *be* insertion error. In which case, it is possible that the learners use passive forms for the unaccusative verbs because of their semantic similarity with passive verbs. However, Choi (2010) shows that the difference of overpassivization ratio between unaccusative verbs and unergative verbs is not significantly different for Korean young learners of English (see also Nam, 2012). In

addition, Choi shows that many *be* insertion errors occur in typical transitive environments in which the subject is an agent and the object is a theme or patient. Therefore, it seems to be obvious that Korean learners' *be* insertion errors are not limited to overpassivization. Although there may exist further possible causes, in this paper, I will investigate the possibilities within the two influential hypotheses: topic-marker hypothesis and underdeveloped functional category hypothesis.

3. Method

3.1 Research questions

In order to reveal the cause of the *be* insertion error made by Korean learners, I begin with the following two research questions:

(8) Research questions
    a. Is the erroneously inserted *be* a topic marker?
    b. Is the erroneously inserted *be* an underdeveloped verbal
       functional category?

To answer the research questions, I will check the following specific possibilities:

(9) Possibilities to be checked
    a. Are there sentences which belong to NP+be+NP+VP type structure?
    b. Which one is inflected between the *be* verb and the main verb?
    c. Does the *be* verb occur in front of a modal verb?
    d. Does the *be* verb behave as an auxiliary?

The checking points given in (9) can reveal which view is more plausible as the cause of Korean learners' *be* insertion error. If learners frequently use 'NP+be+NP+VP' style sentences as in (10), it would support the topic marker view (Shin, 2001).

(10) a. He is, friend is many.
      b. ku-nun   chinkwu-ka   mahn-ta.
         he-TOP   friend-NOM   be+many-PRES-Dec
         'As for him, he has many friends.'

The sentence in (10b) is the typical Korean double nominative construction which well represents the characteristics of Korean as a topic-oriented language. Therefore, the existence of sentences as in (10a) can support the topic marker view in that the sentence in (10a) is exactly parallel to the sentence in (10b). On the contrary, the absence of such data will support the other views such as underdeveloped verbal functional category view.

The distribution of inflectional morphemes can give some hints about which view has an edge in explaining the cause of the *be* insertion error.

(11) a. I was break the vase.
      b. I is broke the vase.

In (11a) it is the *be*-form that is inflected for tense and agreement. On the other hand, the *be*-form in (11b) is not inflected but the following main verb is. The former type of error data would support the verbal functional category view whereas the latter would support the topic marker view.

The checking point in (9c) involves the following test types of sentences.

(12) a. I am don't study English.
      b. I am will study.

The existence of modal verbs blocks the possibility that the co-occuring *be*-forms are analyzed as inflectional functional categories. Particularly, the occurrence of *do* support as in (12a) eliminates the necessity of additional inflectional categories. As a result, sentences as in (12) would weaken the verbal functional category view. In turn, the existence of the *be*-form in addition to a modal verb will support the topic marker view.

I can also check whether the verb carries out the auxiliary function. Auxiliaries can precede a negative word and undergo inversion. If such a function is identified,

the *be* verb must be an auxiliary, that is, a verbal functional category.

3.2 Data and procedures

To answer the research questions mentioned in the previous section, this study analyzed KELC data. KELC contains six different levels as follows: primary, (PRI), basic(BAS), pre-intermediate (PIN), intermediate (INT), advanced (ADV), and post-advanced (PAD). Research procedures are as follows: first, by using Concord in Wordsmith Tools 5.0, I extracted 8,318 sentences containing *be*-forms in KELC. By checking the sentences manually, I sorted out 1,235 sentences with the erroneous *be* verbs.

4. Result

4.1 Topic comment structure

I first checked if the extracted error data contain topic-comment sentences having NP+be+NP+VP type structure. Among 1,235 error sentences, I was able to identify 26 sentences which has the type at issue. Some of the data are given in (13).

(13)   a. Today is my family go to pizza hot. [084MPINT06]
        b. Helen is family go to the mountain. [129FBAST04]

The sentences in (13) have the typical topic-comment structure in that they are parallel to the Korean double nominative construction in (14).

(14)   a. onul-un   nay kacok-i   pizza hut-ey   kassta
         today-TOP   my family-NOM   puzza hut-LOC   went
         'As for today, my family went to Pizza Hut.'
       b. Helen-un   kacok-i   san-ulo   kassta
         Helen-TOP   family-NOM   mountain-DIR went
         'As for Helen, her family went to a mountain.'

In addition to the typical topic-comment structure, 12 sentences as in (15a) could also count as the topic-comment structure.

(15)   a. Because, now time is does not live in stadium. [188MPINT07]
        b. mother is my sister dressing help. [108MBAST07]

Different from the typical topic-comment sentences in (13), the sentence in (15a) does not have the overt subject NP after the *be*-form. The initial NP is not the semantic subject of the main verb and the omitted subject can be understood from the previous context. Therefore, it seems to be obvious that the sentence in (15a) belongs to the topic-comment construction like those in (13). On the other hand, the sentence in (15b) is suspicious in that the NP after the *be*-form is the object of the main verb. 8 sentences belong to this type, and all of this type of errors were made by basic level learners. One possible explanation for this type of errors is that learners not only use *be* verbs erroneously, but also adopt Korean word order so that they put the object in front of the main verb.

If I exclude the wrong word order sentences, only 38 sentences are sorted out as the type at issue, and it accounts for about 3 percent of the whole error data. It is obvious that 3 percent is not enough to confirm the topic marker view. However, given that the topic-comment structure is not prevalent in Korean either, I believe that these data can count as supporting evidence for the topic marker view.

4.2 The distribution of inflection

Through the corpus investigation, I was able to identify 1,128 error sentences that have the 'be + main verb' sequence. The Table 1 illustrates the types of 'be + main verb' sequences which account for about 91 percent of the whole error data at issue.

Table 1. Distribution of inflection: *be*+main verb sequences

Structure	N	%
be + V (present)	807	65
be + V (past)	127	10
be (past) + V (present)	110	9
be (past) + V (past)	84	7
Total	1,128	91

As Shown in Table 1, the *be* + base form sequence accounts for 65% of the whole data. This type of sequence includes the following subtypes:

(16)   a. Car is stop. [395Mintt09]
         b. The strange people is come to the there. [369fpint09]

Sentences of this type do not have tense although most of the sentences describe a situation that happened in the past. In addition, 768 sentences in this type contain a *be* form that agrees with the subject as shown in (16a). The *be* verbs of the other 39 sentences do not agree with the subject.[1]

Although this type of sentences do not have the *be* forms with tense, they can be classified as an underdeveloped inflectional category. This is because most of them still deliver the agreement information. However, as argued by Ahn (2003), it is also possible that the topic marker inflects for agreement, but not for tense. Since both approaches have their own explanation for this type of errors, it is not easily determined which approach is more compatible with the data.

There are also 110 sentences that contain the past forms of *be* when the following main verbs have present forms.[2]

(17)   a. I was eat Pizza [206FINTT06]
         b. they were go pizza shop in 3 o'clock [080MADVT06]

Examples of this type can be strong evidence for the underdeveloped functional

---

[1] All of the sentences, except 8 sentences, contain *is* in them. Given that *is* is the most frequently used be-form, they can be simple mistakes. The other 8 sentences contain *are* in them.

[2] Among the sentences, 4 sentences contain main verbs which agree with the subject as in *sea was looks very cool*.

category approach. This is because the typical role of the inflectional category is to deliver the information of tense. That is, the *be*-forms can be considered the lexical realization of the tense and agreement information. However, this type of errors account for only 9% of the whole error data, and therefore they are not conclusive data to make the decision in favor of the underdeveloped functional category analysis.

On the other hand, there are 127 cases where the present form of verbs precedes the past tense verbs.

(18)   a. famous singer is came in the restaurant. [395MINT09]
   b. her parents are went on a beach. [092MADVT04]

As suggested by Ahn (2003), when the main verb operates normally in terms of inflection, the necessity of the additional functional category disappears. Therefore, sentences of this type can serve as a piece of counterevidence against the functional category approach. However, before taking a position in favor of the topic marker analysis, it must still be explained why the *be*-forms still conjugate in terms of agreement (Yang, 2001, 2002).[3] Regarding this problem, Ahn (2003) suggests that the topic NP and the following *be*-form constitute a constituent and that they establish an internal concord relation. Although both considerations are resonable, they do not have empirical evidence. Therefore, I will also have to reserve the decision for this type of data although they still appear to support the topic marker view.

Finally, there are also 84 cases where both the *be*-forms and the following main verbs appear in past forms, as given in (19).

(19)   my family was bought a pizza and coke. [120FADVT06]

This type of data account for 7% of the whole data and appears to support the functional category analysis in that the *be*-forms inflect for tense. If the *be*-form is a topic marker, there will be no reason to assume that it inflects for tense. Thus, the

---

3   In fact, there was only 1 sentence in which the *be*-form does not agree with the subject: *Jully and Masa is sat in same chair*. Therefore, it seems to be maintained that the *be*-form agrees with the subject regardless of the tense information.

be-forms in past tense support the functional category analysis. Then, the past form of the main verbs will be explained by the wrong tense concord between the *be*-forms and the following main verbs.

To sum up, the distribution of inflection in my data allows both possibilities. That is, the *be* (present) + V (present) construction that is the majority of the data opens both options for the explanation. The *be* (past) + V (present) or *be* (past) + V (past) constructions support the functional category analysis whereas the *be* + V (past) construction supports the topic marker view. Since both analyses can be compatible with certain types of inflectional distribution, I cannot assert which option is more appropriate. However, instead, it will be possible to suggest that both analyses have their own empirical bases.

4.3 The *be*-forms preceding a modal verb

The corpus investigation using KELC extracted 23 sentences which contain the *be*-forms preceding a modal verb as given in (19).

(20)   a. father is doesn't like pizza or cola. [369FPINT09]
         b. me and other people are must be careful. [049FBAST01]

Among the 23 sentences, 20 sentences contain a negative form of verbs such as *don't, didn't, doesn't, wasn't, isn't* and *can not*. The other 3 sentences contain after the *be*-forms modal verbs such as *will* and *must*.

In general, the modal auxiliary verbs are considered fully featured finite verbs. Since finite verbs are assumed to be responsible for the subject licensing, the *be*-forms preceding the finite verbs cannot be additional inflectional functional categories (Chomsky, 1982). In this sense, the existence of *be*-forms in front of the modal verbs supports the topic marker view although this type of examples just accounts for 1.9% of the whole error data.

4.4 The *be*-forms following a modal verb

There are 8 cases in which the *be* form follows a modal verb.

(21) I think ship will be go fast. [054MINTT05]

The *be*-form cannot be considered a topic marker in that it appears after a modal verb. That is, the *be*-form is positioned one word away from the potential topic. Since a topic marker can never be separated by a word from the topic, the *be*-form must not be a topic marker.

On the other hand, it is also suspicious if the *be*-form is a functional category. Since it follows a modal verb, which is a fully featured functional category, I am not sure what kind of function the *be*-form carries out. However, given that the functional category I am dealing with is defected, the *be*-form is more likely defined as a functional category rather than a topic marker. In any case, the number of errors of this type is too small to decide the plausibility of the two possibilities.

4.5 The *be*-forms as an auxiliary verb

There are several criteria that reveal the auxiliary status of a verb in English. For instance, if a verb is contracted with a negation or precedes a negation, it must be an auxiliary verb.

(22) a. monkey isn't eat bananas. [082MPINT02]
      b. I think her is not want this picnic. [018MINTT07]

The corpus investigation extracted only 16 cases of this type.

Another type where the *be*-form is considered an auxiliary verb is shown in (23).

(23) a. Why are you call Pizza hut? [049FINTT06]
      b. Are you find a dog? [370MPINT09]

In (23) the *be*-forms are inverted with their subjects. Since inversion is a characteristic of English auxiliary verbs, the *be*-forms in (23) are also defined as auxiliary verbs. These inverted *be*-forms were observed in 10 cases. Combining the two types of auxiliary *be*-forms, I identified 26 cases, which accounts for about 2% of the whole error data. Since the auxiliary verb is one of the representative functional categories, these data seem to be counted as supporting evidence for the

functional category approach.

5. Conclusion

There can be various different reasons for Korean learners' excessive use of the *be*-forms. In this paper, focusing on the two prevalent hypotheses, I carried out a corpus study using KELC to examine which view explains the phenomenon more plausibly. I first extracted 8,318 sentences containing *be*-forms. By checking all the sentences manually, I sorted out 1,235 sentences which contain erroneous *be* verbs. In this corpus study, I examined the four possibilities which may support the topic marker view or the underdeveloped functional category view. The result shows that neither of them solely explains the error data.

My corpus study shows that learners often use the *be*-form without the purpose of delivering inflectional information. For example, *be*-forms occur after an actual topic but before a subject. This structure is parallel to Korean focus-topic construction. In addition, there are some cases in which the *be*-forms occur before a modal verb. Since modal verbs are considered functional categories with full-fledged inflectional features, the *be*-forms should not be an inflectional category, but something else, for example, a topic marker.

I also checked the distribution of the tense and agreement in my error data. In most cases, the *be*-forms change correctly in accordance with the agreement feature of the subjects. However, there are cases where the tense is carried by only the main verb. There were also cases in which the tense was carried by the *be*-form. These cases support the topic marker view and the underdeveloped functional category view, respectively. The underdeveloped functional category analysis can be supported by some cases in which the *be*-forms follow a modal verb. In addition, the cases where the *be*-forms are considered auxiliary verbs also count as supporting evidence for the view.

The corpus study shows that Korean learners' *be* insertion errors are not explained by a sole reason. Instead, I suggest that the errors involve both language transfer effect such as topic-marker analysis and underdeveloped functional categories shown in learners' developmental stages. The truth can be that there are further reasons that I could not identify in this paper. While focusing on the two popular

hypotheses, I did not pursue an exhaustive investigation into the reasons of the errors. However, when I consider how prevalent the mistake is among Korean learners of English, it must not be a task that can be ignored. I will leave this task for my future study.

References

Ahn, Sung-Ho. 2003. A Note on the Topic-Comment stage in Korean EFL Syntactic Development. *Studies in Generative Grammar* 13(2): 369-382.

Choi, Incheol. 2010. Second language acquisition of English causative alternation: Focusing on overgeneralization errors. *Studies in Modern Grammar* 59: 197-204.

Chomsky, Noam. 1982. Some concepts and consequences of the theory of government and binding. MIT Press.

Hahn, Hye-Ryeong. 2000. *UG Availability to Korean EFL learners: A longitudinal study of different age groups.* Doctoral Dissertation, Seoul National University, Seoul.

Ionin, T. and 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.

Kim, Kitaek. 2011. Overgenerated *be* from Topic Marker to Verbal Inflection. *In selected Proceedings of the 2009 Second Language Research Forum,* 70-81.

Levin, B. and M. Rappaport Hovav. 1995. *Unaccusativity: at The Syntax-lexical Semantics Interface.* Cambridge, MA: MIT Press.

Li, Charles N. and Sandra A. Thompson. 1976. *Subject and topic: a new typology of languages,* 458-489. New York: Academic Press.

Nam, Hyojung. 2012. *Be Insertion Errors by Korean EFL Leaners.* Master Thesis. Kyungpook National University, Daegu.

Shachter, Jacquelyn and William Rutherford. 1979. Discourse function and language transfer. *Working Papers on Bilingualism* 19: 1-12

Selinker, Larry. 1972. 'Interlanguage'. *International Review of Applied Linguistics* 10: 209-310.

Shin, Jung Sun. 2000. *Functional category acquisition by Korean EFL learners: The role of UG in foreign language learning.* Doctoral Dissertation. Seoul National University, Seoul.

Shin, Jung Sun. 2001. L1 influence in foreign language learning: Topic-prominence in Korean EFL learners' interlanguage grammar. *Foreign Language Education* 8: 1-21.

Yang, Hyun-Kwon. 2001. Categorial properties of Korean EFL learners' "Be". *Korean*

Journal of English Language and Linguistics 1: 59-80.

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

Yip, Virginia. 1995. *Interlanguage and Learnability. From Chinese to English.* Amsterdam /Philadelpia: John Benjamins.

Zobl, Helmut. 1989. Canonical typological structures and ergativity in English L2 acquisition. In S. Gass and J. Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition,* 203-221. Cambridge: Cambridge University Pres

Incheol Choi

Department of English Education
Kyungpook National University
80 Daehakro, Bukgu
Daegu 702-701, Korea
E-mail: incheol@knu.ac.kr