KOREAN LEARNERS’ *BE*-INSERTION ERRORS: A LEARNER CORPUS-BASED APPROACH

Incheol Choi
incheol@Knu.ac.kr
Kyungpook National University
Overpassivization errors

- Second language researchers as well as ESL teachers have noted that ESL learners make *be*-insertion errors.
  - Error examples
    - a) *The most memorable experience of my life was happened 15 years ago.
    - b) *Most of people are fallen in love and marry with somebody.
    - c) *My mother was died when I was just a baby.
      (Zobl 1989)
  - Hubbard (1944): overwhelming majority of them appear with intransitive verbs known as unaccusatives (see also Hubbard and Hix, 1988; Yip, 1995; Hwang, 1997; Han, 19987; Ju, 2000)
Korean learners’ *be* insertion errors

- Korean learners of English tend to use the verb *be* excessively not only in the unaccusative environments, but also in other contexts.

  a. *The girl is very very like dog.*
  b. *He is eat hamburger.*
  c. *Bird was flew into the sky.*

  (Choi 2013)
Previous approach I

- Overpassivization (Chung 2010; Zobl 1989; Yip 1995; Lee & Choi 2011)

  - The similarity of the underlying structures between the unaccusative verbs and the passive verbs explains Korean learners’ *be* insertion errors.

- John arrived: \([IP \ e past \ [VP \ arrive \ John]]\)
- John was blamed: \([IP \ e past \ [VP \ blame \ John]]\)
Previous approach I

- Adversity effect (Masuko 1996; cf. Shibatani 1990)
  - Learners do not wrongly passivize intransitive verbs such as sleep and walk. They tend to passivize intransitive verbs when the activities denoted by the verbs affect their internal arguments adversely.
  - occur, happen, and suffer
  - They were suffered by the treatment.
  - kare wa tsuma ni sin are -da
  he Top wife by die Pass Past
  ‘He was adversely affected by his wife’s death.’ (Masuko 1996)
Previous approach II

- topic marker view” (Hahn, 2000; Shin, 2000; Ahn, 2003)

  - The erroneously inserted *be* is a topic marker.

  - *kim-un apeci-ka puca-i-si-ta*
    - Kim-TOP father-NOM rich-COP-DECL
    - ‘As for Kim, his father is rich.’

  - A language transfer effect
    - He is friend is many “As for him, he has many friends” (Shin 2001)
Previous approach III

• functional category view” (Yang, 2001).
  • an underdeveloped verbal functional category: the carrier of tense and agreement
  • Sally is eat bread.
    ‘Sally eats bread.’
Aims of the presentation

• To show the overall pattern of Korean learners’ be insertion errors.

• To reveal possible reasons for the mistakes.

• To reveal the developmental characteristics of Korean learners’ be insertion errors.
A Learner Corpus Study: KELC

• KELC: KNU English Learner Corpus
  • 168,526 words; 830 texts written by 232 children;
  • 8 times essay writing tests over 2 years in the children’s English program in KNU.
  • Annotated corpus: part of speech on the basis of BNC coding system
  • 6 subcorpora in accordance with the proficiency levels: Placement tests: (Bae and Lee 2012)
    • Primary: 1,435
    • Basic: 24,012
    • Pre-intermediate: 39,454
    • Intermediate: 41,128
    • Advanced: 31,210
    • Post-advanced: 31,287
KELC: Be insertion errors

- be errors were retrieved from KELC through manual checking by researchers.

<table>
<thead>
<tr>
<th>Types</th>
<th>PRI</th>
<th>BAS</th>
<th>PIN</th>
<th>INT</th>
<th>ADV</th>
<th>PAV</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>V Tokens (A)</td>
<td>309</td>
<td>4,637</td>
<td>6,701</td>
<td>8,686</td>
<td>5,135</td>
<td>4,700</td>
<td>30,168</td>
</tr>
<tr>
<td>be verbs (B)</td>
<td>131</td>
<td>1,401</td>
<td>1,866</td>
<td>2,042</td>
<td>1,155</td>
<td>1,000</td>
<td>7,595</td>
</tr>
<tr>
<td>be errors (C)</td>
<td>25</td>
<td>331</td>
<td>374</td>
<td>279</td>
<td>176</td>
<td>33</td>
<td>1,217</td>
</tr>
<tr>
<td>C / A * 100</td>
<td>8.1</td>
<td>7.1</td>
<td>5.6</td>
<td>3.2</td>
<td>3.4</td>
<td>0.7</td>
<td>4</td>
</tr>
<tr>
<td>C / B * 100</td>
<td>19.1</td>
<td>23.6</td>
<td>20</td>
<td>13.7</td>
<td>15.2</td>
<td>3.3</td>
<td>16</td>
</tr>
</tbody>
</table>
Research Questions

• Are the *be* verbs a passive auxiliary?
  • checking the main verbs that co-occur with the erroneous *be* verbs

• Are the erroneously inserted *be* verbs a topic marker?
  • NP (topic) + *be* + NP + VP
  • NP + *be* + Aux + VP

• Are the *be* verbs an underdeveloped functional category?
  • checking the distribution of verbal inflection
### Verbs occurring together with *be*-forms

<table>
<thead>
<tr>
<th>verb</th>
<th>Freq</th>
<th>verb</th>
<th>Freq</th>
<th>verb</th>
<th>Freq</th>
<th>Verb</th>
<th>Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>go</td>
<td>215</td>
<td>say</td>
<td>20</td>
<td>sit</td>
<td>10</td>
<td>fall</td>
<td>7</td>
</tr>
<tr>
<td>eat</td>
<td>97</td>
<td>like</td>
<td>16</td>
<td>call</td>
<td>9</td>
<td>find</td>
<td>7</td>
</tr>
<tr>
<td>play</td>
<td>61</td>
<td>sleep</td>
<td>15</td>
<td>climb</td>
<td>9</td>
<td>get</td>
<td>7</td>
</tr>
<tr>
<td>see</td>
<td>50</td>
<td>Smile</td>
<td>15</td>
<td>write</td>
<td>9</td>
<td>stop</td>
<td>7</td>
</tr>
<tr>
<td>look</td>
<td>47</td>
<td>Run</td>
<td>13</td>
<td>catch</td>
<td>8</td>
<td>surprise</td>
<td>7</td>
</tr>
<tr>
<td>come</td>
<td>43</td>
<td>will</td>
<td>13</td>
<td>cry</td>
<td>8</td>
<td>bring</td>
<td>6</td>
</tr>
<tr>
<td>draw</td>
<td>31</td>
<td>ride</td>
<td>12</td>
<td>fly</td>
<td>8</td>
<td>drink</td>
<td>6</td>
</tr>
<tr>
<td>make</td>
<td>30</td>
<td>start</td>
<td>12</td>
<td>help</td>
<td>8</td>
<td>live</td>
<td>6</td>
</tr>
<tr>
<td>finish</td>
<td>29</td>
<td>take</td>
<td>12</td>
<td>talk</td>
<td>8</td>
<td>love</td>
<td>6</td>
</tr>
<tr>
<td>sing</td>
<td>27</td>
<td>know</td>
<td>11</td>
<td>walk</td>
<td>8</td>
<td>rest</td>
<td>6</td>
</tr>
<tr>
<td>have</td>
<td>26</td>
<td>think</td>
<td>11</td>
<td>arrive</td>
<td>7</td>
<td>buy</td>
<td>5</td>
</tr>
<tr>
<td>give</td>
<td>20</td>
<td>laugh</td>
<td>10</td>
<td>disappear</td>
<td>7</td>
<td>drive</td>
<td>5</td>
</tr>
</tbody>
</table>
Result: overpassivization

• Among the 51 lexical verbs, only 5 verbs are categorized as legitimate unaccusative verbs (Levin 1993): go, come, arrive, disappear, and fall.
• The unaccusative verbs: 22.9% (279 cases)
• The other: 77.1% (938 cases)
• go and come are in a gray area (Dowty 1979).
  a. Family are go to the sea. [KELC: BAS]
  b. Friends is go to the picnic. [KELC: BAS]
• arrive, disappear, and fall: 1.7% (21 cases)
Adversity effect

- Sentences involving adversity implication
  - I know that my nose were disappear. [KELC: INT]
  - but dog ‘Funny’ was disappeared. [KELC: PAD]
  - She is fall down on the floor. [KELC: INT]
  - if one guy are fall down and very sad. [KELC: INT]
- Surely *disappear*, and *fall* induce more be-insertion errors than the verb *arrive* or other non-accusative verbs: adversity effect.
- However, the effect is extremely limited.
Result: Topic marker

- NP+ be + (NP) + VP: 39 cases
  - Today is my family go to pizza hot. [KELC: PINT]
  - Helen is family go to the mountain. [KELC: FBAS]
  - now time is does not live in stadium. [KELC: PINT]

<table>
<thead>
<tr>
<th></th>
<th>PRI</th>
<th>BAS</th>
<th>PIN</th>
<th>INT</th>
<th>ADV</th>
<th>PAV</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>0</td>
<td>17</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>Error rate (%)</td>
<td>0</td>
<td>5.1</td>
<td>3.2</td>
<td>2.1</td>
<td>1.6</td>
<td>2.9</td>
<td>3.2</td>
</tr>
</tbody>
</table>
Result: Topic marker

- Ahn 2006
  - In the past computers is are tool of calculate machine. Computers have many advantage. Above all, speed is most advantage. Speed as well as, memory's ability is same millions of people.

- Only one case in KELC
  - 70% well. all is are good but robot was killed [KELC: INT]

- Other four cases
  - Sea was looks very cool [KELC: INT]
  - Anny were goes to beach [KELC: PIN]
Result: topic marker

- Other 125 cases
  - famous singer is came in the restaurant [KELC: INT]
  - He is killed the bird. [KELC: INT]
  - Friends is touched a chicken [KELC: PINT] (the only case of disagreement)
Result: topic marker

• be + Aux + VP: 24 cases
  • be + modal (must, will): 6
    • me and other people are must be careful. [092MADVT04]
    • My class is will go to summer camp. [083MPINT01]
  • be + Negation (don’t, doesn’t, didn’t, wasn’t, can not): 18
    • father is doesn’t like pizza [369FPINT09]
    • police is don’t know. [018MINTT06]
Result: Underdeveloped functional category

• The distribution of inflection
  • Type A: The strange people is come to the there. [KELC: PIN]
  • Type B: Jenny mother is made Gimbab. [KELC: PIN]
  • Type C: I was eat Pizza [KELC: INT]
**Result: Underdeveloped functional category**

<table>
<thead>
<tr>
<th></th>
<th>pri</th>
<th>bas</th>
<th>pin</th>
<th>int</th>
<th>adv</th>
<th>pad</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>total be errors</strong></td>
<td>25</td>
<td>331</td>
<td>373</td>
<td>279</td>
<td>176</td>
<td>33</td>
<td>1,217</td>
</tr>
<tr>
<td><strong>be[prs] + V (base)</strong></td>
<td>25</td>
<td>253</td>
<td>251</td>
<td>161</td>
<td>91</td>
<td>18</td>
<td>799</td>
</tr>
<tr>
<td></td>
<td>(100)</td>
<td>(76.4)</td>
<td>(67.3)</td>
<td>(57.7)</td>
<td>(51.7)</td>
<td>(54.5)</td>
<td>(65.7)</td>
</tr>
<tr>
<td><strong>be[prs] + V (past)</strong></td>
<td>0</td>
<td>37</td>
<td>38</td>
<td>23</td>
<td>24</td>
<td>3</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>(11.2)</td>
<td>(10.2)</td>
<td>(8.2)</td>
<td>(13.6)</td>
<td>(9.1)</td>
<td>(10.3)</td>
<td></td>
</tr>
<tr>
<td><strong>be[past] + V</strong></td>
<td>0</td>
<td>7</td>
<td>52</td>
<td>71</td>
<td>51</td>
<td>8</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td>(2.1)</td>
<td>(13.9)</td>
<td>(25.4)</td>
<td>(29)</td>
<td>(24.2)</td>
<td>(15.5)</td>
<td></td>
</tr>
</tbody>
</table>
Result: Underdeveloped functional category

• *be*-forms after a modal verb: 9 cases
  • my father will be buy toy for me [KELC: INT]
Result: Underdeveloped functional category

- *be*-forms as an auxiliary verb
  - Negation
    - monkey isn’t eat bananas. [KELC: PINT]
    - I think her is not want this picnic. [KELC: INT]
  - Inversion
    - Why are you call Pizza hut? [KELC: INT]
    - Are you find a dog? [KELC: PIN]

<table>
<thead>
<tr>
<th></th>
<th>pri</th>
<th>bas</th>
<th>pin</th>
<th>int</th>
<th>adv</th>
<th>pad</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aux + be + V</td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>be + SUB + V</td>
<td></td>
<td></td>
<td>4</td>
<td>5</td>
<td>1</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>be +not +V</td>
<td></td>
<td></td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>14</td>
</tr>
</tbody>
</table>
Discussion

• Conspiracy effect

• Language transfer: topic marker effect, adversity effect, …

• Underdeveloped verbal functional category: Once the functional projection is established (e.g., INFL), it tends to be filled with a be-form. (cf. Hawkins 2001)

• Whatever the reason (e.g. topic marker), it seems to function as an inflectional category eventually… (agreement, tense, inversion, negation)
A question

• Then, why do particularly Korean learners make more be-insertion errors?
  
  • cf. According to Hubbard (1983)’s observation from nonnative learners’ English compositions, more than 90% of be-insertion errors takes place with unaccusative intransitive verbs. Therefore, the majority of the errors were identified as the overpassivization errors.

  • As we observed, Korean learners’ be-insertion errors cannot only be attributed to the overpassivization errors: conspiracy effect (topic, adversity effect, …)
A possible factor

- *be* may function as the general sign of the beginning of the VP (or lexical realization of INFL).
  - *be* verbs / total verbs
    - BNC: 3,042,613/13,200,285 = 23.05%
  - Elementary text book (Chunjae)

<table>
<thead>
<tr>
<th></th>
<th>3-4 year</th>
<th>5-6 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total V</td>
<td>2,477</td>
<td>2,704</td>
</tr>
<tr>
<td>Total be</td>
<td>1,043</td>
<td>876</td>
</tr>
<tr>
<td>rate</td>
<td>42.1</td>
<td>32.4</td>
</tr>
</tbody>
</table>

- KELC

<table>
<thead>
<tr>
<th></th>
<th>PRI</th>
<th>BAS</th>
<th>PIN</th>
<th>INT</th>
<th>ADV</th>
<th>PAD</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be</td>
<td>131</td>
<td>1,401</td>
<td>1,866</td>
<td>2,042</td>
<td>1,155</td>
<td>1,000</td>
<td>7595</td>
</tr>
<tr>
<td>V</td>
<td>309</td>
<td>4,637</td>
<td>6,701</td>
<td>8,686</td>
<td>5,135</td>
<td>4,700</td>
<td>30,168</td>
</tr>
<tr>
<td>Rate(%)</td>
<td>42.4</td>
<td>30.2</td>
<td>27.8</td>
<td>23.5</td>
<td>22.5</td>
<td>21.3</td>
<td>25.2</td>
</tr>
</tbody>
</table>