

Grammatical and ungrammatical uses of intransitive verbs in essays written by Japanese learners of English: A large-scale corpus analysis

Kazuharu Owada¹, Hajime Tsubaki², Eiichiro Tsutsui³ and Victoria Muehleisen⁴

¹Tokyo College of Music, ²GITI, Waseda University, ³ Faculty of Pharmaceutical Sciences, Hiroshima International University, ⁴ School of International Liberal Studies

owadakazu@aoni.waseda.jp, hjm.tsubaki@asagi.waseda.jp, tsutsui@ic.hirokoku-u.ac.jp,
vicky@waseda.jp

Abstract

There have been many studies on the acquisition of English intransitive verbs in the field of learner corpora. Most of these studies have so far concluded that even advanced learners of English ungrammatically passivize intransitive verbs and produce sentences such as ‘*the accident was happened’ and ‘*the mobile phone was appeared.’ These ungrammatical passives thus produced are often referred to as ‘passive unaccusatives’ by many researchers.

Oshita (2000) is probably the first to analyze these ungrammatical passives of intransitive verbs in a written learner corpus. He used the Japanese data consisting of 1,363 essays from the Longman Learners Corpus (LLC). However, in order to gain a more thorough understanding of Japanese learners of English, an analysis of a much larger corpus is needed.

The purpose of this study is to examine ungrammatical passivization of three intransitive verbs by using a large-scale learner corpus.

Keywords

learner corpus, intransitive verbs

1 Introduction

In this study we examine how frequently three intransitive (unaccusative) verbs (*appear*, *happen*, and *occur*) are ungrammatically passivised through a university-level essay corpus called the SILS-JLE Corpus, which is a sub-corpus of the SILS English Learner Corpus, compiled at the School of International Studies at Waseda University.

First, we will briefly discuss the content of the SILS-JLE Corpus. Then we will examine the frequency of ungrammatical passivization of three target intransitive verbs: *appear*, *happen*, and *occur*.

And lastly, we will conclude by mentioning several limitations of this study.

2 The method

2.1 SILS-JEL Corpus

We have so far conducted several studies on the English learners’ use of verb patterns using a learner corpus called the ‘SILS English Learner Corpus.’ This is a corpus compiled at the School of International Studies at Waseda University which contains more than 5,600 essays by college-level students. This corpus includes information such as the nationality of students, the draft number, the class level and the TOEFL scores.

Because of the massive amount of data the SILS English Learner Corpus contains, we focused our attention on the first draft essays written by the Japanese students and named this sub-corpus ‘SILS Japanese English Learner Corpus (henceforth, SILS-JEL Corpus).’

2.2 Vocabulary in the SILS-JEL Corpus

To determine the total number of words of this sub-corpus and its vocabulary level breakdown, we ran the JACET 8000 program to see what levels of indexes (types) and tokens (total words) are used in the first-draft essays by the Japanese university students. The JACET 8000 consists of eight levels, level 1 being the most basic 1,000 words and level 8 being the most difficult 1,000 words.

Table 1 shows the breakdown of indexes and tokens used in the essays. The total words of this sub-corpus turned out to be 1,997,701. Words in level 1 constitute about 80% in terms of tokens, and make up about 5% in terms of indexes. As for the frequency in terms of indexes, words that fall outside level 8 are most frequently used, making up about 58% of the total indexes.

Table 1: Words used in the SILS-JEL Corpus

	indexes	%	tokens	%
level 1	1114	4.977	1601488	80.167
level 2	991	4.427	127737	6.394
level 3	939	4.195	53302	2.668
level 4	892	3.985	50524	2.529
level 5	789	3.525	19457	0.974
level 6	729	3.257	13428	0.672
level 7	633	2.828	12376	0.62
level 8	592	2.645	8797	0.44
over 8	13144	58.723	84413	4.226
cont. forms	765	3.418	10534	0.527
proper nouns	1794	8.015	15644	0.783
non-words	1	0.004	1	0
total	22383	100	1997701	100

3 The study

3.1 Purpose

Although the three verbs *happen*, *occur*, and *appear* fall in the level 1 of JACET 8000, and are thus are considered to be basic words that require a full mastery, previous studies have shown that even advanced English learners tend to passivize these three verbs, producing ungrammatical sentences. In this study, we investigated the frequency of such ungrammatical uses of *happen*, *occur*, and *appear* in the SILS-JEL Corpus.

Since this is a preliminary investigation, we restricted ourselves to the word sequence of the target words using the *Kwic Finder* software. With this method, we were not able to take into consideration the occurrence of adverbials in the sequences.

3.2 Results and discussion

The frequencies of ungrammatical uses of *appear*, *happen*, and *occur* were 33, 20, and 33, respectively, as can be seen in Tables 2-4.

Table 2: Ungrammatical uses of *appear* in SILS-JEL Corpus

<i>appear</i>	class_level			
sequence	Advanced	Inter.	Basic	Total
be appeared	6	1	1	8
am appeared	0	0	0	0
is appeared	4	2	1	7
are appeared	8	2	0	10
was appeared	4	2	0	6
were appeared	1	0	0	1
has been appeared	1	0	0	1
have been appeared	0	0	0	0
had been appeared	0	0	0	0
Total	24	7	2	33

Table 3: Ungrammatical uses of *happen* in SILS-JEL Corpus

<i>happen</i>	class_level			
sequence	Advanced	Inter.	Basic	Total
be happened	3	4	1	8
am happened	0	0	0	0
is happened	5	5	1	11
are happened	1	0	0	1
was happened	0	0	0	0
were happened	0	0	0	0
has been happened	0	0	0	0
have been happened	0	0	0	0
had been happened	0	0	0	0
Total	9	9	2	20

Table 4: Ungrammatical uses of *occur* in SILS-JEL Corpus

<i>occur</i>	class_level			
sequence	Advanced	Inter.	Basic	Total
be occurred	4	2	1	7
am occurred	0	0	0	0
is occurred	1	5	1	7
are occurred	0	0	0	0
was occurred	2	3	0	5
were occurred	1	1	2	4
have been occurred	0	0	0	0
had been occurred	0	0	0	0
Total	8	11	4	23

The results show that Japanese English learners do produce ungrammatical passive uses of *appear*, *happen* and *occur*, and interestingly, such production does not seem to vary much by class level. That is, students in the advanced level classes as well as in the intermediate and basic level classes produced these ungrammatical forms.

4 Conclusion and limitations

We have observed that Japanese English learners produced ungrammatical intransitive verbs (i.e., passive unaccusative verbs) in their essays.

However, this study has several limitations. First, we need to incorporate the ratio of grammatical uses of the target verbs as opposed to the ungrammatical uses. Second, we need to thoroughly investigate the occurrence of adverbials. The adverbials that might have occurred in the sequence of the target verbs were deliberately left out due to the heavy processing load of the corpus. To overcome this problem, tagging of the corpus is definitely needed for further analysis.

Acknowledgement

This work was supported by JSPS KAKENHI 22520585.

References

Oshita, H. (2000). What is happened may not be what appears to be happening: a corpus study of 'passive' unaccusatives in L2 English. *Second Language Research* 16, 293–324.